

Powerful visions for astronomy

US astronomers are getting better at forming a collective view of priorities. Their latest wish list is worth fighting for, but it is the process used to generate it that gives the community the best chance of attracting the necessary funds.

More than 100 US astronomers have put their heads together and come up with a prioritized wish list for public investment in their field during the next decade (see page 381). This 'decadal review' of astronomy and astrophysics, produced under the auspices of the National Academy of Sciences, is to be applauded. It is one of science's most effective priority-setting exercises: a model for astronomers in other regions to follow and for other disciplines to consider.

The process is all-important. Washington is drowning in reports, and the last thing congressional staffers or White House budget examiners want is to settle differences among squabbling scientists. So a coherent approach within the community is essential for making others sit up and listen.

This has necessitated a year of full-time leadership by some key individuals. Each of the nine sub-panels — organized mostly by observational technique — held three meetings. There were also 'town meetings' and public sessions to gauge the community's views. Each sub-panel voted on priorities for its category, ranking projects primarily on their scientific usefulness.

These recommendations then went up to the main committee, which held five meetings over an 11-month period. All agree it was hard to make choices. Each of the sub-panel reports went through the full academy review process, and the 23 reviewers of the main report had credentials equal to anyone on the committee.

They generated 120 pages of comments — a painful process, but all part of shoring up the report's credibility. And the knowledge that this is a chance to feed one's ideas into such a well-regulated and influential process ensures that some of the best of the community's creative vision gets captured.

The plan presented last week contains plenty of exciting science. An 8-metre infrared telescope in space could look back to the dawn of the Universe to see the first galaxies that formed. Its ground-based

companion, a 30-metre segmented telescope, would see proto-planetary disks as sharply defined rings instead of blurry blobs. The Terrestrial Planet Finder — which the committee would like to see get under way at least in the next decade — would be capable of finding the first Earth-like worlds.

The astronomers who served on the review committee believe that, in the coming decade, astronomy can lead the way in stimulating the public's interest in science. But the community will have to show that it is being frugal with public money. That leaves little room for duplicative projects or those born of institutional or national pride. The review panel is right to make a strong case for cooperation, not only among US astronomers and their international colleagues, but also between privately and publicly owned observatories.

The decadal review committee's plan would go a long way towards achieving that goal. We could look forward to having a powerful array of instruments scanning the Universe across the electromagnetic spectrum. Many of these projects are already planned to be international, which, like the survey itself, reflects a welcome maturity in the astronomical community.

The report also emphasizes the need to follow through with the recommendations of the last review while getting these new projects under way. The Atacama Large Millimeter Array, although approved, is the only one of that committee's four major recommendations that is not yet under construction. It must not fall by the wayside.

The presentation of this work is what matters now. There is going to be a glossy version of the report, published later this year, aimed at the general public. As for the leading astronomers who compiled the report, they will be astronomy's ambassadors. They've already met with the National Science Foundation and the space agency NASA, and will brief Congressional staffers and the Office of Management and Budget. As hard as they've laboured in the past year, in some important ways their work has only just begun. ■

Values of the abstract

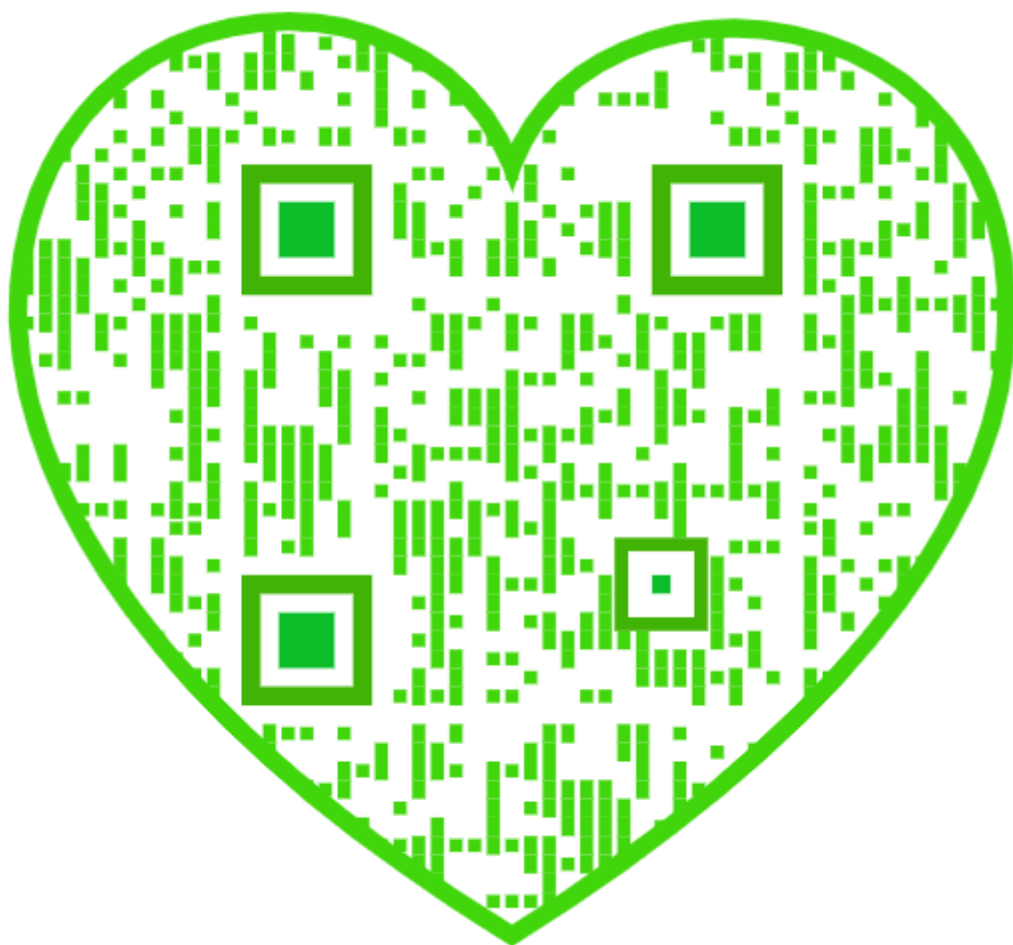
A new set of prizes is an apt celebration of the significance and wonder to be found in pure mathematics.

"I have just one chance for escaping a verdict of complete triviality, that I may be judged to have created something worth creating. And that I have created something is undeniable: the question is about its value." As the great number theorist G. H. Hardy went on to conclude in *A Mathematician's Apology*, his work, producing nothing that was useful, had simply "added something to knowledge, and helped others to add more; and these somethings have a value which differs in degree only, and not in kind, from that of the creations of the great mathematicians, or of any other artists, great or small, who have left some kind of memorial behind them."

Science will grab anything that it can turn to its own ends, and plenty of seemingly useless mathematics has turned out to be essential for future breakthroughs. When Hardy wrote his book in 1940, physicists hadn't discovered the possible links between the Riemann

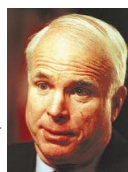
hypothesis on the distribution of prime numbers and the statistics of eigenstates of quantum systems with underlying classically chaotic behaviour. And science poses fundamental mathematical challenges of its own, such as the unsolved Navier–Stokes equations of fluid dynamics and the Yang–Mills equations derived as a geometrical account of the fundamental forces between elementary particles.

All credit, then, to the Clay Mathematics Institute for celebrating the intrinsic value of such challenges. The institute has gone beyond the Fields Medal and the Wolf Prize in singling out long-standing mathematical problems — the three above, plus four others that are less obviously applicable — and valuing each of their solutions at US\$1 million (see page 383). It's an excellent way for a private foundation to recognize the eternal fascination that mathematics holds for people such as Hardy, and for the rest of us. ■





Star crazy
Britain makes moves
to join European
Southern Observatory
p382



Heated debate
US senator shows
signs of taking global
warming seriously
p385



HIV latest
Plans for *Salmonella*
bacterium to deliver
oral AIDS vaccine
p386



Danger man
Prince Charles stirs
controversy over
risks of science
p388

US astronomers draw up their wish list for a decade of funding

Tony Reichhardt, Washington

Bigger, better and a bit more expensive: that's the next generation of astronomical observatories envisaged by a panel of more than 100 scientists who last week laid out a ten-year road-map for US astronomy and astrophysics.

The planning exercise, organized once every decade by the National Academy of Sciences, is closely watched by government agencies that fund large astronomy facilities.

The 'decadal review' for 2000–10 was co-chaired by Christopher McKee, a theoretical astrophysicist at the University of California, Berkeley, and Joseph Taylor of Princeton University, who shared a Nobel Prize for finding evidence of gravitational waves from a binary pulsar.

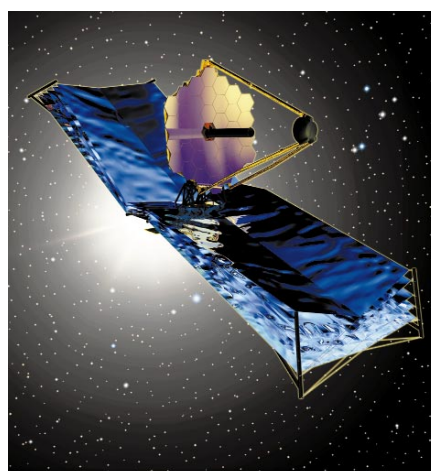
Formed in 1998, McKee and Taylor's committee and nine sub-panels spent nearly a year ranking priorities for the next decade of astronomical observation, both space-based and ground-based.

Their top recommendation is the Next Generation Space Telescope (NGST), a \$1 billion successor to the Hubble Space Telescope. Orbiting a million miles from Earth, NGST will use an 8-metre deployable mirror to observe in the infrared, where it will see the oldest and faintest galaxies with hundreds of times the sensitivity of Hubble. The US space agency NASA, in cooperation with the European Space Agency and Canada, hopes to launch the NGST in 2009.

Second on the wish list is a 30-metre-diameter ground-based telescope known as the Giant Segmented Mirror Telescope (GSMT), which is similar in design to the current Keck telescopes in Hawaii but three times their size.

Just as Keck complements Hubble's high-resolution images with high-resolution spectroscopy, the GSMT and NGST would give optical-infrared astronomers a powerful one-two punch for addressing problems ranging from the evolution of galaxies to the search for planets around other stars.

Several concepts have emerged recently for giant ground-based telescopes with apertures ranging in size from 30 to 100 metres. In recommending an instrument at the low



Top tip: the Next Generation Space Telescope is the advisory panel's number one priority.

end of that scale, the panel chose caution over performance.

"When we looked at the technology development for a 100-metre telescope, we got scared," says Taylor. Not that the GSMT

couldn't eventually be reconciled with the 100-metre 'Overwhelmingly Large Telescope' (OWL) concept being studied by the European Southern Observatory, he says (see page 382). The panel recommends that half the GSMT's \$350 million cost come from either international or university partners.

Other major initiatives — defined as space facilities costing over \$500 million and ground-based projects over \$50 million — are, in order of priority: the Constellation-X Observatory, a suite of four space-based telescopes that would succeed the current Chandra X-ray observatory and focus its investigations on black holes; an upgrade to the Very Large Array (VLA) of radio telescopes in New Mexico; and the Large-Aperture Synoptic Survey Telescope (LSST), a wide-field, fast-read-out telescope that could survey the entire visible sky down to faint objects at the 24th magnitude every week.

Formerly known as the 'Dark Matter Telescope', the LSST was renamed to better reflect the variety of subjects it will address, from cataloguing supernovae to finding

France sets up elite Internet school

Heather McCabe, Paris

France is to create a *grande école* — an elite higher-education institution — dedicated to the Internet, French Prime Minister Lionel Jospin announced last week. It will begin teaching courses in fields from computer science to electronic commerce next year.

The 'Grande École de l'Internet' will be near Marseilles — joining a 'technopole' of existing technology-orientated businesses and institutions. Its running costs, estimated at about FF60 million (US\$8.2 million) each year, will be shared by the government, the University of Aix-Marseille, two *grandes écoles* organizations — the Groupe des Écoles des Télécommunications (GET) and Les Écoles des Mines — and private industry.

One of the new institution's tasks is to redress the country's shortage of computer

engineers — projected to reach 60,000 by 2005, according to Claude Guéguen, the scientific director of GET. Jospin's research, technology and space adviser, Bertrand Mabilie, says: "It's a real problem to find enough people to work in the industry, especially in start-up technology companies."

The two-year programme will be open to students who have completed master's-level work at engineering schools and universities, or who have completed their first year at a *grande école* of engineering.

Gilles Kahn, scientific director of INRIA, France's computing research agency, welcomes the new *grande école*, but says much more is needed.

"We can't content ourselves with creating this one institution," he says. "We must make efforts across the board to improve technology education."

- ▶ 90 per cent of all near-Earth asteroids larger than 300 metres across.

The panel's highest-priority 'moderate' mission is not the most expensive on the list, but many astronomers consider the proposed \$50 million Telescope System Instrumentation Program to be vital to the health of their field. By funding new instruments for private telescopes such as the Keck, which is owned by California universities, the federal government will be buying a certain amount of access for all US astronomers.

Other moderate initiatives include the Gamma-ray Large Area Space Telescope (GLAST), which will succeed the Compton Gamma Ray Observatory; a pioneering gravity-wave telescope in space called LISA (for Laser Interferometer Space Antenna); an Advanced Solar Telescope; and — a rarity for astronomy — an instrument to be placed on the International Space Station, known as the Energetic X-ray Imaging Survey Telescope (EXIST).

The top 'small' initiative identified in the report is the National Virtual Observatory (NVO), a programme to compile astronomical data gathered in different wavelengths into a single large database accessible to scientists and the public. John Bahcall, an astrophysicist with the Princeton Institute for Advanced Study who led the last decadal review in 1991, likens the NVO to a hot "growth stock" on Wall Street. "Many of the most important discoveries will come from mining this database," he says.

The overall programme proposed by the McKee–Taylor panel would cost \$4.7 billion up to 2010 — 20 per cent more than the wish list offered by Bahcall's committee ten years

Prioritized equipment initiatives and estimated federal costs, 2000–10

Initiative	Cost (\$ million)
Major initiatives	
Next Generation Space Telescope (NGST)	1,000
Giant Segmented Mirror Telescope (GSMT)	350
Constellation-X Observatory	800
Expanded Very Large Array (EVLA)	140
Large-Aperture Synoptic Survey Telescope (LSST)	170
Terrestrial Planet Finder (TPF)	200
Single-Aperture Far InfraRed (SAFIR) Observatory	100
Moderate initiatives	
Telescope System Instrumentation Program (TSIP)	50
Gamma-ray Large Area Space Telescope (GLAST)	300
Laser Interferometer Space Antenna (LISA)	250
Advanced Solar Telescope (AST)	60
Square Kilometre Array (SKA) technology development	22
Solar Dynamics Observer (SDO)	300
Combined Array for Research in Millimetre-wave Astronomy (CARMA)	11
Energetic X-ray Imaging Survey Telescope (EXIST)	150
Very Energetic Radiation Imaging Telescope Array System (VERITAS)	35
Advanced Radio Interferometry between Space and Earth (ARISE)	350
Frequency Agile Solar Radiotelescope (FASR)	26
South Pole Submillimetre Telescope (SPST)	50
Small initiatives	
National Virtual Observatory (NVO)	60
Other small initiatives	246
Total for all initiatives	4,670

ago. Early in its deliberations, the committee asked NASA and the National Science Foundation, the two funding agencies responsible for most US astronomy, how fiscally conservative it should be, and were told "not to restrain ourselves", says Taylor.

If past decadal reviews are any guide, the new report should be well received on Capitol Hill and in the White House budget office. All four major initiatives proposed by Bahcall's committee in 1991 have since been approved, as have nine out of 11 moderate-size projects. Indeed, the astronomy surveys are almost unique in their influence on Washington decision-makers.

William Wulf, president of the National Academy of Engineering, believes the reason is that they "are not typical please-send-money reports — they represent some tough decisions".

Indeed, says Bahcall, the exercise is a "heart-wrenching" process in which astronomers "have to turn down proposals that were a decade in the making".

The committee also tried to consider input from the wider astronomical community, not just the élite, according to Lee Anne Willson, an astronomer at Iowa State University who chairs the astronomy section of the American Association for the Advancement of Science.

She set up a web page for rank-and-file astronomers to provide comments to the survey committee, and says that the comments were read. The committee's final decisions were also thoroughly vetted — after the sub-panels and panels did their work, another 23 readers weighed in with opinions as part of the National Research Council's review process.

The report probably comes as close to a consensus as it could, given the time it took to produce, says Willson. "Astronomy should be proud of itself."

▶ <http://books.nap.edu/catalog/9839.html>

Britain asks to join European Southern Observatory

David Dickson, London & Alison Abbott, Munich

Keen to see the creation of a single European body with the clout to raise the funds needed for future large telescopes, Britain is this week opening negotiations about possible membership of the European Southern Observatory (ESO).

Based in Munich, the eight-member ESO is responsible for the construction and operation of a number of telescopes in the Southern Hemisphere, notably the Very Large Telescope (VLT) currently being built on Mount Paranal in Chile.

Britain declined to join the organization when it was set up in 1962, claiming that it could not afford to do so. But Ian Halliday, the chief executive of the Particle Physics and Astronomy Research Council (PPARC), said last week that are now several reasons for becoming a member.

As well as giving British astronomers access to the VLT, membership of the ESO



Halliday: ready to negotiate over fees.

would allow the PPARC to become directly involved in the planning of major new international facilities, such as the proposed optical telescope known as the Overwhelmingly

Large Telescope (OWL).

"In 10 years' time, if we want to build this type of facility, we need to be part of an organization such as ESO," says Halliday. "We will not be able to afford to build it ourselves."

Many members of the ESO, including its director Catherine Cesarsky, are happy for Britain to join. Franco Pacini, a member of the ESO council and director of the Arcetri

Astrophysics Observatory in Florence, says: "No-one would deny how important it will be for ESO to have Britain as a member. It is a pity that [Britain] did not join long ago."

But agreeing on the terms of entry will not be easy. The treaty setting up the ESO stipulates that new members must pay an entry fee as a contribution to the capital costs of the facilities they will be able to use.

In Britain's case, the entry fee would be £55 million (US\$81.8 million). But Halliday says that, given the pressures on the UK science budget, paying such a sum is out of the question. "If they are insistent about us paying our full whack, we will not be able to get into ESO," he says.

One ESO official says: "It is an unfortunate moment [for Britain to ask for special treatment] because ESO has just made a huge investment in the VLT, and member states may not be too happy about ▶

► Britain joining without making a retrospective contribution.”

Cesarsky says that the ESO will enter negotiations with a flexible outlook. An offer in kind — for example, giving astronomers from ESO member states access to British telescopes at Paranal — would be considered seriously. But to suggest that the ESO would reduce the total value of the entry fee is, she says, “pure dreaming”.

Even if the question of the entry fee is settled, there is also the issue of the annual payments, which would currently amount to £12 million a year. Without a significant increase in government funding to cover this sum, there would have to be cuts in other parts of the PPARC's budget.

Some of the money could be found, for example, by closing some of Britain's existing telescopes on La Palma in the Canary Islands. This prospect is worrying astronomers working on the facilities. There is also concern from particle physicists that covering a commitment to the ESO could be achieved by taking money away from accelerator research.

But the importance to British astronomers of joining the ESO was underlined in a report, *Unveiling the Universe*, published by the PPARC last week. It lists joining the ESO as high among its priorities. Reflecting the priorities of the PPARC's astronomy panel, the report suggests that the research council should seek an extra £39 million from the government over two years.

Halliday is not only confident of success in his negotiations with the ESO, but also optimistic that his efforts to build closer European collaboration will be received sympathetically in Whitehall. “ESO is currently not a well funded organization,” he says. “They probably have more telescopes than they can afford. Somewhere in there is a potential for negotiation.”

However, Cesarsky says that the ESO is not so strapped for cash that it needs to compromise to suit Britain's pocket. “Britain knows what the conditions are for joining,” she says. ■

Mathematicians chase the seven million-dollar proofs

David Dickson

Mathematics might be the queen of the sciences, but as far as international prizes go, its omission from the list of Nobel disciplines has made it a poor relation.

All that, however, is about to change. The recently established Clay Mathematics Institute (CMI) in Cambridge, Massachusetts, was due to announce this week that it is offering \$1 million for the proof of each of seven classical mathematical problems.

Among them is the Riemann hypothesis concerning the distribution of prime numbers — widely considered the most important problem in number theory — and the Navier–Stokes equations familiar to anyone who has studied fluid dynamics.

Also listed is the ‘P vs NP’ (polynomial versus non-deterministic polynomial) problem. This concerns the relationship between the difficulty of solving a problem and of checking solutions, and bears directly on the theoretical possibility of being able to break computer codes (see *Nature* 400, 115–116; 1999).

“We wanted to make a statement to the general public that maths is important,” says Arthur Jaffé, professor of mathematical physics at Harvard University and a member of the CMI's Science Advisory Board. “We believe that the future of the world is going to be dependent on science, and that mathematics is at the centre.”

Jaffé was largely responsible for persuading businessman Landon T. Clay, 18 months ago, to set up a non-profit foundation dedicated to increasing and disseminating mathematical knowledge. The CMI already funds a range of activities, from research fellowships to workshops and summer schools. The institute now hopes that its decision to offer a series of prizes will recreate the type of public fervour that accompanied the solution of Fermat's last theorem in 1993 by the British-born mathematician Andrew Wiles.

“We wanted to do something original,” says Alain Connes, professor of analysis and geometry at the Collège de France in Paris, which hosted the meeting at which the prizes were to be announced. Like Wiles, Connes is a member of the CMI's four-member Science Advisory Board.

It is almost a century since German mathematician David Hilbert set out 23 key unresolved problems that he predicted would set the agenda for mathematics in the new century, during a famous lecture in Paris, on 8 August 1900. His prediction was accurate. The problems he set out dominated much of mathematics in the past 100 years, and all but three have been resolved. (The solution of



the Riemann hypothesis is one of the three outstanding problems from Hilbert's list.)

In contrast to Hilbert's speech — and reflecting concern in some quarters that the CMI might be attempting a similar task — those responsible for the prize emphasize that they have no desire to determine the direction in which mathematics moves forward.

Rather, they are trying to recreate excitement about the activity of mathematics, both in the general public and among school students. “Sometimes people have the wrong idea of maths and think that it will be overtaken by computers,” says Connes. “The seven problems, each selected by top specialists in their respective fields, are totally inaccessible to computers.”

The competition has strict rules for determining whether a problem has actually been solved. In particular, a proposed solution must be published in a refereed journal “of worldwide repute”, and must have gained “general acceptance” in the mathematics community within two years of publication. Also, to ensure that the lure of a reward does not encourage mathematicians to lock themselves away and work on the problems in private, prior work will be taken into account in awarding a prize.

There is no time limit to the competition. “We know that they are difficult, because people have been working on them for a long time,” says Jaffé. But he also expresses optimism: “I would hope to live to see each of these solved.”

The \$1 million problems are:

- The P vs NP problem
- The Riemann hypothesis
- The Poincaré conjecture
- The Hodge conjecture
- The Birch and Swinnerton-Dyer conjecture
- Navier–Stokes equations
- Yang–Mills theory

► <http://www.claymath.org>

ESO



Counting the cost: will Britain pay the price for access to Europe's Very Large Telescope?

BioMed Central boosted by editorial board

Declan Butler

The idea that the primary scientific literature should be available online free of charge gained heavyweight support this week, when ten prominent biomedical scientists agreed to join the editorial board that advises on BioMed Central's (BMC) overall goals.

The new publishing house, which is owned by the Current Science Group (see *Nature* **402**, 110; 1999), started accepting papers this week. It will peer review and publish free on the web papers across the biomedical sciences.

The ten scientists include Elizabeth Blackburn of the University of California, San Francisco, Steven Hyman, director of the US National Institute of Mental Health, Marc Kirschner, head of the Department of Cell Biology at Harvard Medical School, Philippe Kourilsky, director-general of the Institut Pasteur in Paris, Paul Nurse, director-general of the Imperial Cancer Research Fund in London, Harold Varmus, president of the Memorial Sloan-Kettering Cancer Center, and Mitsuhiro Yanagida, president of the Molecular Biology Society of Japan and Kyoto University.

BioMed Central will accept papers in around 50 disciplines, such as neuroscience, virology and developmental biology. As well as referees, each area will have a series of coordinating scientists.

Reviewers will assess scientific accuracy,

not interest, says Harvey Shoolman, BMC's editorial coordinator. For example, he says, it will publish negative results and details of single experiments.

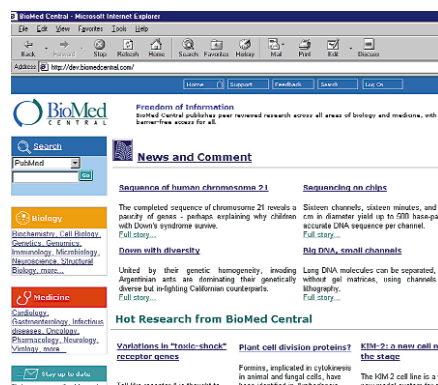
Accepted papers will be accessible at <http://www.biomedcentral.com> and archived on PubMed Central (PMC), the free repository for life-science papers launched earlier this year by the US National Institutes of Health. Copyright will remain with authors. If BMC takes off, it will give a boost to PMC, which at the moment contains material from only a handful of journals.

"Our central task will be to persuade authors to publish with us," says Vitek Tracz, chief executive officer of the Current Science Group. "We don't know how big or small the problem will be."

He adds that the group will use technology to offer carrots such as accelerated publication and efficient searching. It will also recruit scientists to highlight important papers as they are published.

Editorial board member Joseph Martin, dean of the faculty of medicine at Harvard Medical School, is one of many who argues that scientists are unlikely to shift *en masse* from print to electronic journals. He says this is precisely why a strong signal from leaders of the biomedical establishment was needed.

Hyman admits that "BioMed Central is not the only way to skin a cat". But he says he agreed to join the board to send a signal that many biomedical researchers believe that the



BioMed Central: will provide free access to peer-reviewed research.

current system of journals is obsolete, and that the Internet offers improved alternatives.

Like many scientists, both Martin and Hyman say they read only a handful of the top journals. For the rest, their main need is to be able to search the primary literature without running up against the subscription firewalls of thousands of journals.

"If I cannot access a paper freely I just ignore it," says Hyman, adding that, although he "respects the peer review provided by the existing journals system", he feels that this could be maintained within an electronic database of the primary literature.

Both men argue that the bulk of the biomedical literature is 'balkanized' across thousands of low-impact journals, and that scientists would be better served by a seamless, searchable and freely accessible database. "We pay for the science, why should the journals get copyright?" says Hyman, arguing that the Internet makes the cost of publication derisory compared with the cost of producing the research in the first place.

One big question is how BMC is going to pay for the services it will provide. Shoolman says it can afford to run at a loss for some time thanks to Tracz's personal fortune. For his part, Tracz says that "the business model is unknown to everyone, including me, but we are confident we have enough potential sources of revenues".

Tracz believes his venture is the forerunner of a market where the primary literature is free, and where commercial publishers provide other services, such as reviewing, reporting and alerting. One of BMC's major activities will be selling database services in partnership with other companies and groups of scientists, he adds.

"Easy access stimulates usage, and barriers reduce usage," says Andrew Odlyzko of AT&T Research Laboratories in Florham Park, New Jersey, who studies electronic publishing. He says that scientists will soon find that unless their papers are freely available, they might as well not be written.

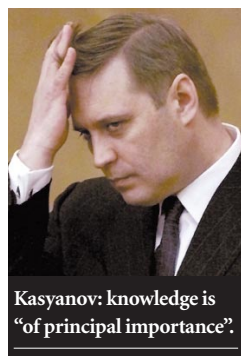
Russian PM supports science

Carl Levitin, Moscow

Russian science appears to have a friend in the country's new prime minister, Mikhail Kasyanov — at least in principle. In a major address to the lower house of the Russian parliament last week, Kasyanov said: "We now have entered the information epoch" where knowledge is "of principal importance".

He went on to say that "the development of education and promoting basic and applied research in the sphere of modern technologies should be considered as the most important investment in our country's future". But despite this show of concern, the new government will operate without the former Ministry of Science and Technologies.

Kasyanov has reduced the overall number of ministries from 30 to 24. Science now comes under a single Ministry of Industry, Science and Technologies. Headed by Aleksandr Dondukov, 46, the new ministry will be responsible for many aspects of Russia's economy, including its industrial development, military production



and arms exports. Many fear that basic science will not be high on its list of priorities.

Dondukov graduated from the Moscow aviation university in 1957, and worked on aircraft design in several Soviet and later Russian design bureaux. In 1994 he was appointed chairman of the directors' council of the Yakovlev Design Bureau in Moscow.

His stated ambition was to centralize control of aircraft production in Russia by working closely with the aviation research and development offices of the former Soviet Union republics. He is not known to have previously expressed any interest in basic research or educational programmes.

and arms exports. Many fear that basic science will not be high on its list of priorities.

Dondukov graduated from the Moscow aviation university in 1957, and worked on aircraft design in several Soviet

and later Russian design bureaux. In 1994 he was appointed chairman of the directors' council of the Yakovlev Design Bureau in Moscow.

His stated ambition was to centralize control of aircraft production in Russia by working closely with the aviation research and development offices of the former Soviet Union republics. He is not known to have previously expressed any interest in basic research or educational programmes.

Congress wakes up to climate change

Colin Macilwain, Washington

Advocates of action to tackle climate change won an influential ally last week, as Senator John McCain's (Republican, Arizona) Senate Commerce Committee heard leading scientists testify that humankind is causing significant global warming.

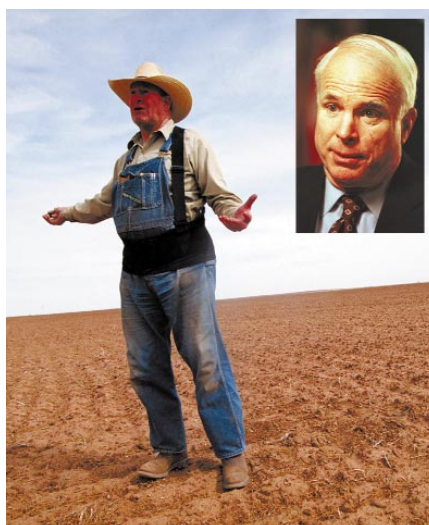
Other congressional hearings in recent years have pitched global warming 'sceptics' against the larger group of climate scientists who believe that greenhouse-gas emissions are contributing to climate change — creating the impression that researchers are in disarray. But McCain's hearing presented a more sober picture of the evidence.

Neal Lane, Bill Clinton's scientific adviser, said there was now "a significant consensus" on the issue. "Basically, the debate has changed from 'are we warming the Earth?' to 'how much are we warming the Earth?'" Lane said.

Each of the five scientific witnesses called before the hearing repeated this perspective. These included John Christy, a climate scientist at the University of Alabama, who has sometimes been sceptical about the human component of climate change.

McCain, who recently failed to win the Republican nomination for this November's presidential election, said the hearing was a response to concerns encountered on the campaign trail.

"In town-hall meeting after town-hall



Warm words: in the wake of a US drought, John McCain (inset) speaks out on climate change.

meeting, young people would stand up and say: 'What is your position on global warming?'" McCain said. "I don't have a plan, but I do believe that policy-makers should be concerned about mounting evidence of global warming."

Environmentalists greeted his approach as further evidence that the tide of US opinion is slowly turning in favour of action to limit greenhouse-gas emissions.

In the past year, major corporations

including Ford, General Motors and DuPont have withdrawn their support from the Global Climate Coalition, which has campaigned against such action. Some of these corporations have also set their own internal targets for reducing emissions.

Until McCain's hearing on 17 May, Congress had not reflected this shift. It is trying to prevent the US government spending money on carbon-emission trading schemes or other components of the Kyoto Protocol. The Clinton administration, which signed the protocol, has not asked the Senate to ratify it, knowing it would refuse.

Drafts of the third assessment report of the Intergovernmental Panel on Climate Change, to be published next year, state that human activities are contributing to global warming.

The US government, meanwhile, is set to publish its own national assessment of the consequences of climate change next month. This is intended to influence the debate by showing Americans the relevance of the problem, but some industrial groups have already questioned its neutrality.

The next conference of the parties to the Kyoto Protocol will take place in The Hague in November, a week after a presidential election that will help to determine the United States' position. Whatever happens, advocates of the Kyoto Protocol do not envisage the US Senate ratifying it before 2003. ■

French unions upset by plans for physiology centre

Heather McCabe, Paris

France's biomedical research agency, INSERM, has clashed with some of its research staff over an ambitious proposal to build a new research centre outside Paris. Focusing on physiology and physiopathology, the centre would cost FF500 million (US\$69 million). It would have five animal facilities, holding almost 20,000 animals, and would be equipped with advanced imagery and computer systems.

The aim of the proposal is to collect the research activities of several institutions under one roof. It would bring together physiologists currently working for INSERM, the National Centre for Scientific Research (CNRS), the National Institute for Agronomic Research (INRA), and the life-sciences department of the French Atomic Energy Commission (CEA). Researchers from the Paris-Sud XI University in Orsay and the University of Paris XII in Créteil would also be involved.

INSERM officials say the centre would provide a boost to research in physiology,

which has suffered over recent years compared with the more popular areas of the life sciences, such as molecular biology.

But while accepting the need to redress this balance, labour unions representing research staff argue that the money would be better spent on strengthening and upgrading existing research centres.

Plans for the project are awaiting the green light from the Ministry of Research. One author of the proposals, Marc Peschanski, head of an INSERM research unit on neuroplasticity and therapeutics in Paris, says he expects a response within the next few months.

The regional council of the Ile-de-France, which covers Paris, has already voiced its support for the proposal. According to Peschanski, the region has agreed in principle to finance half of the project over two years, although it will not provide any funds before 2002.

The new centre also has the support of Claude Griscelli, the director-general of INSERM, who describes the project as "very

interesting". But he adds: "We should not put everything in one institute. We need to put resources into existing facilities. One cannot be done without the other."

The National Union of Scientific Research Workers (SNTRS) — one of the unions representing scientific employees — goes further. It says that, before creating a single large research centre, government research agencies should first build on existing physiology centres to form a modern network of facilities.

Union member Jean-Pierre Bazin of INSERM's laboratory on quantitative medical imagery in Paris says that the proposal for the 'physiopol' "lacks the sort of long-term scientific planning necessary for a project of this size". He believes that the scientific objectives have not been well identified.

Bazin points out that scientists working for the other institutions to be included in the centre have not yet been consulted. Peschanski admits that consultations have yet to take place, but says the project is only in the preliminary stage of planning. ■

Bacterial AIDS vaccine ready for testing

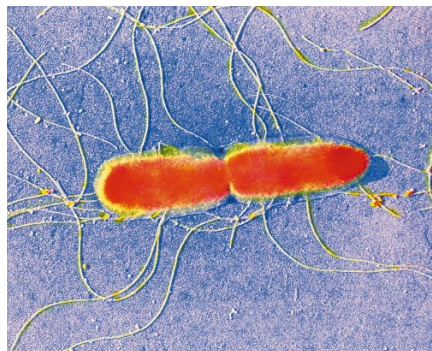
Paul Smaglik, Baltimore

Robert Gallo, director of the Institute of Human Virology in Baltimore, Maryland, last week unveiled plans to test a novel AIDS vaccine in Baltimore and Uganda within the next two years. Unlike most vaccines against HIV, it can be delivered orally. Also unusual is its use of a bacterium — a weakened form of *Salmonella typhi* — to deliver a package of HIV genes.

An orally administered vaccine for AIDS would be relatively cheap and easy to administer. Gallo believes the oral *Salmonella* approach may be particularly useful, as it seems to stimulate strong immune responses in mucous membranes, thought to be the key to preventing the heterosexual transmission of HIV. This is the most common way for the disease to spread in Africa and throughout the developing world.

But in announcing the planned trial, Gallo was at pains not to promise too much. This reflects past experience: despite more than a dozen clinical trials for preventative AIDS vaccines, no one candidate has emerged as a clear solution to the pandemic.

The vaccine was developed by George Lewis, director of the Institute for Human



Germ warfare: weakened *Salmonella* could deliver an immune-boosting cargo of HIV genes.

Virology's vaccine division, and his colleague David Hone. It consists of a strain of *Salmonella* that cannot itself cause disease, carrying genes for an HIV protein called gag and fragments of several other viral proteins. Andrew McMichael, of the University of Oxford, originally chose this package for use as a 'naked' DNA vaccine.

Once inside human cells, the cell walls of the weakened *Salmonella* break and release the HIV DNA. The host cell then transcribes this DNA, to produce the HIV proteins

which trigger an immune response. The goal is not to elicit the production of antibodies against HIV, but to prompt cytotoxic T-lymphocytes to recognize and destroy any cells infected with the virus.

The vaccine's development is being sponsored to the tune of US\$1 million over the coming year by the International AIDS Vaccine Initiative (IAVI), which has been given US\$25 million from the Bill and Melinda Gates Foundation. Wayne Koff, IAVI's vice-president for research and development, says the jury is still out on the vaccine: "It's unclear if these are the right genes at this time."

However, one advantage of the *Salmonella* system is that the relatively large bacterium could carry a suite of different genes. Gallo expects to tinker with other possible gene components in the laboratory while this approach is being tested.

Tony Fauci, director of the US National Institute of Allergies and Infectious Diseases, agrees that it pays to be cautious. "We have been burned before in trying to predict how a candidate will fare before the trial even starts," he says. "Having said that, I like this approach."

♦ <http://www.iavi.org/4/index.asp?highlight=134>

Tax blow leaves Australian science reeling

Peter Pockley, Canberra

The Australian government has slipped a cut to research grants into the fine print of its latest budget, delivered on 9 May. Australian researchers have learnt that they are expected to aid the government in making 'savings' of hundreds of millions of dollars.

The impact of a controversial new goods and services tax (GST), being introduced in July, emerged only when organizations received letters from ministers, which *Nature* has obtained, after the budget was presented.

Minister for Industry, Science and Resources, Nick Minchin, wrote that this year the normally reliable budget figures are "indicative only".

The budget plans "will require adjustment to the funding allocation of a large number of programmes", he added. Minchin's department has lost A\$104 million (US\$60 million), possibly because of the new tax, and is moving to pass this on to grantees.

The government argues that financial modelling shows that recipients would have received a 'windfall' from the GST, compared to the other taxes that it will replace.

The letters do not specify the proportion to be cut from each grant. But the news that



Cullen: faces a fiscal "triple whammy".

it could be up to 3.5 per cent set alarm bells ringing in the research community — few, if any, of whom had any inkling of the government's intentions.

The only cut calculated so far is for the

Commonwealth Scientific and Industrial Research Organisation. Its budget would fall 3.4 per cent in real terms, to A\$616 million.

The 65 Cooperative Research Centres (CRCs) are deeply worried about their government grants, which average A\$2 million. Most of their money goes on salaries or is passed through universities, both of which are free of a GST imposition.

Peter Cullen, director of the CRC for Freshwater Ecology in Canberra, says the centres face a "triple whammy" consisting of the GST, the extra costs of administering it, and annual cuts of 1 per cent. He believes that costs for the CRCs are likely to increase with the GST, with no provision for compensation.

The government is not commenting, but the Labor opposition will press ministers for figures when parliamentary committees resume next week. The row has overshadowed the government's increase in support for R&D. A rise in funding for medical research and tax concessions for business R&D are estimated to boost expenditure by 1.1 per cent in real terms, to A\$4.5 billion (see *Nature* 399, 94; 1999).

In charting the year ahead, Treasurer Peter Costello made no mention of R&D, the only new programme being modest support for commercialization of biotechnology (A\$30.5 million over four years). This failed to quell a chorus of complaint from academic and research leaders over the continuing neglect of universities (see *Nature* 382, 659; 1996).

The annual statement on R&D funding revealed a decline in government support to 0.68 per cent of gross domestic product. Australian business spending on R&D ranks seventeenth out of 20 industrialized nations.

The effect of the new GST on grants to university-based researchers remains so unclear that applicants have been required to submit two budgets with and without adjustments for it. ■

Cool response from agribusiness to royal plea for nature

London Britain's future monarch, Prince Charles, last week made a public plea to "restore the balance between the heartfelt reason of instinctive wisdom and the rational insights of scientific analysis". Delivering the last of this year's BBC Reith lectures, he argued that mankind's "duty of stewardship" for the Earth "has become smothered by almost impenetrable layers of scientific rationalism".

Charles criticized the neglect of research into traditional, 'organic' systems of farming in favour of heavy use of chemicals and the development of genetically modified crops. "We should show greater respect for the genius of nature's designs, rigorously tested and refined over millions of years," he said. "This means being careful to use science to understand how nature works, not to change what nature is, as we do when genetic manipulation seeks to transform a process of biological evolution into something completely different."

The Royal Society responded: "If we are going to [live] sustainably in the coming decades, we must pay greater attention to the insights of science". Cropgen, a group recently set up by the agribusiness industry, said: "Demonizing gene technology inevitably risks increasing the number and variety of chemicals used in farming today."

Green agriculture minister drops biotech conference

Genoa Italy's new minister for agriculture has withdrawn the ministry's sponsorship of TEBIO, one of Italy's largest biotechnology conferences and exhibitions. It was one of Alfonso Pecoraro Scanio's first acts in office.

Scanio, a member of the Green party, said it was a response to a statement made by new prime minister Giuliano Amato.

Amato had identified a series of priorities for his government — including restrictions on field trials of transgenic crops, human cloning and DNA patenting — while presenting his programme to the Italian parliament on 28 April.

Despite the move, TEBIO still took place in Genoa this week. Ironically, it was sponsored by the Italian Committee on Biotechnology and Biosafety, which is under the aegis of the prime minister's cabinet.

Nuclear leaders pledge to eliminate arsenals

Washington The five main nuclear powers — Britain, China, France, Russia and the United States — last week signed an "unequivocal undertaking" to "accomplish the total



Prince Charles: "Use science to understand how nature works, not to change what nature is."

elimination of their nuclear arsenals".

The pledge, made at a meeting of parties to the Nuclear Non-Proliferation Treaty in New York, goes further than previous promises by the powers to dispose of their nuclear weapons. It was welcomed by arms-control groups, although it contains no indication of a timetable for disarmament.

The White House sought to play down the significance of the statement, however, denying that it marked a shift in US policy.

The five powers (then including the Soviet Union) originally agreed to work towards nuclear disarmament when they first signed the Non-Proliferation Treaty in 1968.

German university building programme grinds to a halt

Munich No new university building projects will start in Germany next year unless government funding improves, the country's science council warned last week.

The federal government must significantly increase its financing of university building and large equipment in the next four-year funding round, starting in 2001, said the Wissenschaftsrat.

It warned that the position of universities, whose funds for building have been steadily eroded for nearly a decade, was now "extremely insecure" and that research conditions are frequently poor.

The Wissenschaftsrat has a long waiting list of new projects that it considers urgent. These include renovating dilapidated 1970s buildings, extending research institutes in the former East Germany, and starting work on prestigious new projects such as the planned Centre for Biosciences at the University of Cologne.

Ethical divide on US genome project

Washington The US Human Genome Project's Ethical, Legal and Social Implications (ELSI) research programme is to be split into two. One programme will be overseen by the Department of Energy (DoE), which initiated the HGP in the United States. The other will be run by the US National Human Genome Research Institute (NHGRI), which has had a higher profile in the project in recent years.

The split represents diverging ELSI interests by each group, says Elizabeth Thomson, ELSI research programme director at the NHGRI. The DoE's ELSI efforts have in recent years been focused on privacy and public education issues, whereas the NHGRI's have involved increasing emphasis on genetic testing and the role of genomics in clinical settings.

Thomson announced the split at a National Human Genome Advisory Council meeting this week.

Stone to retire after life of highs and lows in space lab

San Diego The director of the Jet Propulsion Laboratory (JPL), Edward Stone, announced last week that he will retire next year when he reaches the age of 65.

Based in Pasadena, California, the JPL is managed for the US space agency NASA by the California Institute of Technology (Caltech). It has 5,000 employees and a budget of \$1.3 billion.

The JPL has seen both stunning successes and failures in space exploration under Stone's leadership. A physicist and astronomer, Stone led the team that sent the Voyager mission to photograph Jupiter, Saturn, Uranus and Neptune in 1977. But last year two JPL-led missions to Mars failed when spacecraft were lost.

Caltech officials said that a search will begin immediately for a new director.

Anderson resigns as Wellcome governor

London The Wellcome Trust announced last week that Roy Anderson, formerly director of the trust's Centre for Epidemiology and Infectious Diseases at the University of Oxford, has resigned as a governor. Following a series of controversies, Anderson had already announced his resignation from the university, and is to take up a post at Imperial College, London (see *Nature* 403, 472 & 404, 214; 2000).

In a short statement, the Wellcome Trust said it wished to acknowledge the significant contribution he had made as a governor, as well as "his very substantial scientific contributions more broadly".

Changing of the guard

Allan Bradley is to head the Sanger Centre, a powerhouse of the Human Genome Project. Trisha Gura asks how he will meet the challenges.

Talking in pure scientific terms, Allan Bradley — due to become director of the Sanger Centre in October — is an obvious choice to lead the largest DNA sequencing centre involved in the Human Genome Project (HGP) into the post-genomic era. As the human genome is completely sequenced and annotated over the next two years, action at the centre, near Cambridge in England, will shift to establishing the function of each constituent gene. Bradley's current lab at the Baylor College of Medicine in Houston, Texas, is a world leader in elucidating the functions of mouse genes using gene-targeting technology.

"He was our candidate of choice because of the really fantastic work he has done on the analysis of gene function in animal systems," says Michael Dexter, director of the Wellcome Trust, which announced his appointment last week (see *Nature* 405, 264; 2000). But as director of the Wellcome-funded centre, Bradley will face political challenges every bit as daunting as the scientific ones.

For anyone familiar with the fractious politics of genomics, one question looms large: how will Bradley handle relations with Celera? Craig Venter claims that his company, based in Rockville, Maryland, will assemble a finished human genome sequence at about the time the HGP publishes its 'draft' sequence later this year. But while the HGP's sequence is immediately and freely available to all, anyone wanting to access Celera's data will have to accept the company's terms and conditions (see *Nature* 404, 324; 2000).

This has driven a wedge between Celera and the public project, with current Sanger director John Sulston leading criticism of the company. As the disagreement between the HGP and Celera became increasingly acrimonious, relations between Sulston and Venter degenerated. After negotiations between the HGP and Celera (designed to

establish a framework for cooperation) broke down, the two men traded insults in public (see *Nature* 404, 117; 2000).

Significantly, Bradley seems to want to build bridges. He says he is committed to the Wellcome Trust's policy on free public access to genome data, but praises Venter's scientific contribution. Like many researchers, Bradley remains unsure if Celera's 'whole-genome shotgun' sequencing method will provide a short cut to the human genome. But, he says, the approach "has changed the way people are thinking and doing things". And having a powerful competitor has energized the HGP, he claims. "Upon reflection, the element of competition certainly has brought the mouse sequencing to a more urgent priority than it would have been otherwise."

Bradley, who has served on the scientific advisory board of several biotech companies, is also seen as more open to industrial collaboration than his predecessor. Earlier this month, Sulston was quoted by *The*

Mouse studies will be key to identifying the function of human genes

Guardian newspaper as saying: "Global capitalism is raping the Earth." Sulston says his concern was not with industry in general, but with Celera's aggressive stance. "It concerned me that any one company might gain a monopoly on the fundamental operation. But that danger is receding now as more and more companies come in," says Sulston, who is proud of his achievements as Sanger director. "The bulk of the human genome is really and truly out in the public domain. I have



Careful: Bradley (left) will need diplomatic skills.

pushed for that very strongly." But he is less bullish about his disagreement with Venter: "I regret that it became so public."

With the shift into functional genomics, Sulston agrees that Bradley is the right person to lead the Sanger Centre. Raw sequencing will still be important: the mouse genome and probably the zebrafish are among several model organisms the centre is expected to target. But the functional studies that are Bradley's forte will assume a growing importance. "He has a proven record for doing that sort of genetics on a large scale," says deputy director Richard Durbin.

Bradley, a youthful 40, spent his graduate and postgraduate years at the University of Cambridge under the tutelage of geneticist Martin Evans. At Baylor, where he is now a Howard Hughes Medical Institute investigator, he has used gene targeting to disrupt specific genes in mouse embryonic stem cells. By inserting these cells into embryos, and breeding from the resulting mice to obtain animals that are pure bred for the genetic modification in question, he has unravelled the functions of legions of mouse genes. Once the mouse sequence is available alongside the human, similar mouse studies will become key to identifying the function of human genes, especially in disease.

Bradley's professional rivals say the Sanger directorship will test his talents to the full — but they're confident that he's up to the challenge. "He is an outstanding scientist," says Mario Capecchi of the University of Utah in Salt Lake City, a pioneer of mouse gene-targeting studies. "He is quiet but energetic, and he is thoughtful. My feeling is that he is certainly capable of doing a good job."

Trisha Gura is a freelance writer in Cleveland, Ohio.



Solid ground: the Sanger is the biggest sequencing centre in the international Human Genome Project.

Deep thoughts

North America is about to come under intense geophysical scrutiny. Rex Dalton explains how the four projects known as EarthScope will advance our understanding of volcanoes, fault systems and earthquakes.

Astronomy has its huge telescopes; high-energy physics its particle accelerators. Even biology, once the bastion of small science, has the Human Genome Project. But until now, scientists studying the interior of the Earth haven't had a flagship endeavour to propel them into the major league of big science.

In the United States, however, that's about to change, thanks to a bold initiative called EarthScope. The goal is to provide a three-dimensional view of the geophysical processes affecting North America, supplying data in real time. EarthScope's backers predict huge scientific dividends. To quote the initiative's website: "Integrated with new and existing geologic information, these data will provide unprecedented opportunities to unravel the structure, evolution and dynamics of the North American continent."

Our understanding of volcanoes, fault systems and earthquakes will be revolutionized, claim EarthScope enthusiasts. "There are huge implications for understanding seismic hazards," says Mark Simons, a geophysicist at the California Institute of Technology in Pasadena.

President Bill Clinton's farewell budget request to Congress, for the fiscal year 2001, includes US\$19.4 million in start-up funding for the first two components of EarthScope. If the initiative is funded in its entirety, the total cost is likely to rise to more than US\$500 million over the next decade. That should be achieved without squeezing the National Science Foundation's (NSF's) budget for research grants, as the money will be requested from the NSF's Major Research Equipment allocation, the US Geological Survey (USGS) and the space agency NASA.

Into the unknown

While most Earth scientists are excited by EarthScope's promise, many are aware that spending huge amounts of money on big science means tackling serious logistical problems — and taking big risks. Some are already worrying about the potential for cost overruns, and the damage to their field's image should any of the high-profile projects fail to achieve their goals. "We are all taking risks on EarthScope," agrees Mark Zoback of Stanford University in California, a principal investigator on one of its component projects. "All the projects are expensive, but they are taking us to where we have never been before."



Shock tactics: the San Andreas fault (seen here crossing the Mojave Desert) will be the focus of much of EarthScope's efforts as the project attempts to unravel the mysteries surrounding seismic activity.

Indeed, Zoback and others argue that Earth scientists must seize the moment. Advances in instrumentation have made EarthScope technically feasible, they say, while the current favourable budgetary climate has provided the chance of funding. They expect the initiative to foster collaborations between geologists, geophysicists and seismologists, strengthening Earth sciences as a whole. "Each of the projects emerged simultaneously from different communities," says Zoback. "The concept of combining them should give us some real breakthroughs."

EarthScope is made up of four projects. The San Andreas Fault Observatory at Depth (SAFOD) is an ambitious scheme to monitor an active region of the infamous fault.

We pussyfoot around on earthquake prediction. This will bring us much closer to understanding the problem.

Researchers plan to drill a borehole to a depth of 4 kilometres in the mountains south of San Francisco. The USArray project will use an array of seismometers to provide three-dimensional images of the structures that lie deep beneath the entire United States.

The Plate Boundary Observatory (PBO), meanwhile, will study the surface deformation caused as the Pacific and North American plates grind slowly past one another. The final element is a satellite fitted with interferometric synthetic aperture radar (InSAR). This will map tectonically active regions of the continent, allowing researchers to study displacements of the Earth's surface before, during and after earthquakes or volcanic eruptions.

SAFOD and the USArray already have the backing of the NSF's governing body, the National Science Board. The board wants to spend US\$75 million on equipment for both projects during a four-year period from 2001. It also envisions a companion fund of US\$62 million to pay for operations and research over the next decade.

With the SAFOD borehole, scientists will be able to place instruments in and around a zone of earthquake generation, giving them a three-dimensional view of seismological activity. The planned location is near Parkfield, where the San Andreas fault spawns

frequent quakes. The last significant Parkfield earthquake was a magnitude 6 event in 1966, and the USGS had predicted that a similar-sized quake would happen by the end of 1993. Since then, it has become clear that earthquake prediction is much more complicated than was once supposed, and seismologists are still waiting.

Shaking off uncertainty

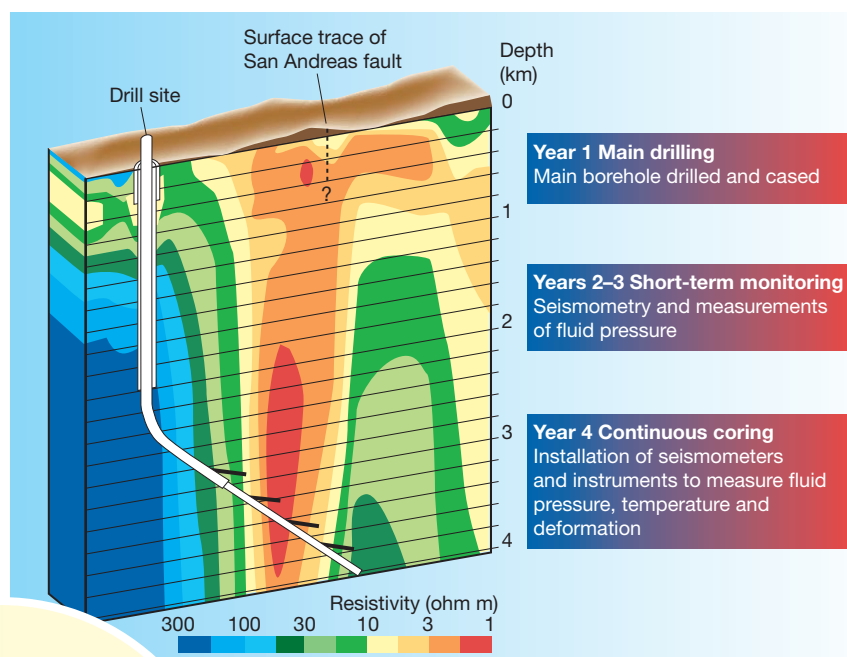
Data from SAFOD should help narrow the uncertainties that make understanding the processes that drive earthquakes, and predicting their occurrence, so difficult. The main borehole will be drilled straight down for about 2 kilometres next to the San Andreas fault, before veering directly through the fault until it reaches a depth of 4 kilometres (see figure, right). From about 3 kilometres down, side holes will be bored into the fault, allowing scientists to extract rock and fluid samples for laboratory tests. Instruments will also be lowered into the shaft to measure the pressure of fluid within pores in the rocks, temperature, stresses within the fault zone, and near-field seismic waves from nearby earthquakes.

These data should address the many unknowns surrounding the physical and chemical processes that affect the San Andreas fault, and provide a firmer understanding of the stresses driving earthquake rupture. Many theoretical models have been put forward to explain the fault's behaviour. For instance, geophysicists have proposed that fluid within the pores of rocks deep within the fault is under high pressure, and that variations in this pressure influence the occurrence of earthquakes. But in the absence of the sort of information that SAFOD should yield, these theories remain untested.

William Ellsworth, a geophysicist with the USGS in Menlo Park, California, and one of the project's coordinators, predicts that the difference between SAFOD's results and those of surface-based projects should be as striking as that between conventional X-rays and the three-dimensional medical images obtained using computerized tomography scanners. "With SAFOD, we can get under the earthquake activity," explains Ellsworth. "This should open up some exciting opportunities for imaging."

Bigfoot steps in

The USArray is a seismic mapping project that will start in California, and will cover the entire nation over a period of 10 years. At its core will be a transportable array of 400 broadband seismometers — dubbed "big-



Going underground: the borehole to investigate seismic processes in the San Andreas will be the first shaft to be drilled so close to a major fault.

foot" by seismologists — that will be placed in a grid spaced at intervals of about 70 kilometres.

In addition, some 2,400 instruments will provide high-density, shorter-term seismic recordings from interesting regions within the larger array. The project will also provide new instruments to upgrade and expand, to more than 100 permanent stations, the National Seismic Network run by the USGS, which monitors areas of high earthquake risk. The NSF is expected to fund the seismometers through the Incorporated Research Institutions for Seismology (IRIS), a consortium of universities which has its headquarters in Washington DC. IRIS already operates a pool of more than 500 seismometers for other geophysical projects.

Both natural seismic events and explosions will be used to draw up a comprehensive picture of the geological structures lying beneath the continental United States. This will include data on the depth of faults, the dimensions of magma chambers under active volcanoes, and the deep structure of sedimentary basins. Through collaborations with seismologists in Canada, and possibly Mexico, the survey could become truly pan-continental (see 'Keeping up with the neighbours', overleaf).

The USArray could image structures such as the deep roots of the North American craton, the ancient rocks that sit beneath the middle of the continent. It will also probe deeper still into the Earth's mantle and core. "A project like this will give us coherent and

continuous images at high resolution — a more holistic view of what is going on in the deeper Earth," says Anne Meltzer of Lehigh University in Bethlehem, Pennsylvania, who is coordinating planning for the USArray.

Defining the daily grind

The PBO has yet to be placed firmly on the NSF's wish-list, but geophysicists regard it as an integral component of EarthScope. The Pacific-North American Plate boundary zone covers a third of the North American continent. Its diverse features include the Rocky Mountains, the Sierra Nevada, the volcanic peaks in the Cascade Mountains and Alaska's Aleutian Islands, and the San Andreas fault system. The PBO will place this zone under unprecedented scrutiny using an interconnected network of global positioning system (GPS) receivers, strain meters, tiltmeters and seismometers.

A backbone of GPS receivers, roughly 100 kilometres apart, will stretch from Alaska to Mexico, and their relative motions will be used to draw up an overall picture of the surface deformation caused by the opposing drifts of the Pacific and North American plates. Higher concentrations of instruments will be placed along the San Andreas fault, and at sites where magma is known to be welling up from deep Earth. These include Yellowstone National Park; Long Valley east of the Sierra Nevada in California; two locations in the Cascades, one of which is likely to be Mount Rainier; and two places in Alaska.

The full PBO network would consist of 1,275 continuously recording GPS receivers and 245 strain meters — of which around 400 GPS receivers and 45 strain meters are already

news feature

in place under existing projects. The strain meters will gauge short-term deformation on timescales from minutes to months, whereas GPS devices work best at charting movement over periods of months and years. Using these instruments, the PBO researchers hope to observe the full range of tectonic activity, including the build up of strain before a quake, and the relaxation that follows.

Paul Silver, a seismologist at the Carnegie Institution of Washington, who is chairing the steering committee for the PBO, says the project offers an unprecedented opportunity to study slow transient deformation — the gradual stop-start movement of the plates over periods of minutes to years. Seismometers miss these movements, Silver notes, recording only fast slips of the plates. “There are many reasons to believe that slow plate movement over a period of time is an important part of the cycle of an earthquake,” he adds. “We pussyfoot around on earthquake prediction. This is an opportunity to explore transient deformation, which will bring us much closer to understanding the earthquake problem.”

Screen shots from the sky

As a natural extension of their effort, Silver and his colleagues want to work with the team planning the InSAR satellite. NASA would be the prime funding agency for the satellite, which is the priciest EarthScope project — costing between US\$200 million and US\$400 million, depending on the type of instruments flown. Although NASA is not yet firmly committed to the project, officials are making encouraging noises. “We are extremely interested in cooperating with the NSF and the USGS,” says Earnest Paylor, the agency’s programme director for solid earth and natural hazards.

The European Space Agency already operates a satellite — ERS-2 — with InSAR technology. Synthetic aperture radar allows precise mapping of the Earth’s topography. InSAR, which analyses sequential images of the same

area, can measure deformation at the Earth’s surface caused by seismic events (see right). But ERS-2 is used for a wide range of different projects. Obtaining its data is costly, and can take months. What’s more, observations of North America aren’t repeated frequently enough for the needs of US geophysicists.

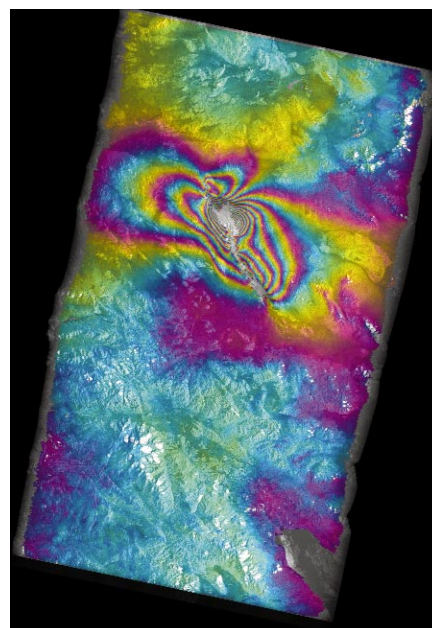
In contrast, the EarthScope InSAR proposal would provide data for free, almost in real time, over the Internet. “Imagine — you come into the lab in the morning, pull up the previous day’s data and watch a volcano deform in Chile, or an earthquake in Taiwan,” says Caltech’s Simons, a member of the InSAR planning team. The satellite’s orbit would be arranged so that it provided images of the region being studied by the PBO every eight days. This would allow scientists to stack radar images to create three-dimensional pictures of earth movement over time, complementing the PBO results.

Digging for victory

While few Earth scientists doubt EarthScope’s scientific potential, some remain nervous about the risks posed by going down the ‘big science’ road. The blow suffered by their colleagues in planetary science, following NASA’s loss of two consecutive Mars probes, is fresh in everyone’s minds.

The biggest questions surround SAFOD, as no one has drilled this close to a major fault before. At the annual meeting of the Seismological Society of America in San Diego in April, the SAFOD team stressed that the precise location for the hole is critical: bore in the wrong place — say, in an old, inactive fault zone — and the budget for the project would be blown with no real scientific returns.

To place the borehole at the best site, the SAFOD team is using magnetotelluric techniques employed by Martyn Unsworth, a geophysicist at the University of Washington in Seattle. Unsworth sends electrical impulses into the Earth using a portable generator, and analyses high-frequency changes



The big picture: this interferometric image of the 1999 Hector Mine quake in the Mojave Desert, California, shows how the ground was displaced.

in electrical and magnetic waves to determine the resistivity of the rocks below. “It’s like viewing a fetus with ultrasound,” says Zoback, one of SAFOD’s coordinators.

Unsworth’s analyses have prompted the planned drilling site to be moved twice. But whatever the location, some scientists fear that shifting geological structures, high temperatures and pressures, or salt-water flows could disrupt the shaft or damage instruments. “This is something that has never been done before,” says Christopher Scholz, a geophysicist at the Lamont-Doherty Earth Observatory of Columbia University in New York. “You can expect massive cost overruns.”

But Zoback rejects such criticisms, pointing to the findings of an NSF review panel that examined the SAFOD plan in 1999 and concluded that it was well designed. However, the panel did question whether a US\$1 million contingency fund set up to cover unexpected problems with the drill project was sufficient.

Viewed in the context of the scientific gains promised by EarthScope, the initiative’s backers argue that the risks associated with SAFOD are entirely justified. Once EarthScope’s results start coming in, they predict, the criticisms will soon be forgotten. “EarthScope will energize Earth science like nothing since plate tectonics,” says Zoback. “We need to move ahead.” ■

Rex Dalton is Nature’s West Coast US correspondent.

EarthScope ♦ <http://www.earthscope.org>

US Geological Survey Parkfield site ♦ <http://quake.wr.usgs.gov/QUAKES/Parkfield/>

Nature debate on earthquake prediction ♦ http://helix.nature.com/debates/earthquake/quake_frameset.html

Incorporated Research Institutions for Seismology ♦ <http://www.iris.edu>

Sample InSAR images ♦ <http://earth.esa.int/INSI/>

Keeping up with the neighbours

In a project that will dovetail neatly with the USArray of seismometers, a Canadian consortium of government, universities and businesses is planning a similar seismic mapping project.

Called POLARIS (Portable Observatories for Lithospheric Analysis and Research Investigating Seismicity), the project would include 90 seismometers and 30 further instruments to measure the resistivity of deep geological structures. It would cost almost Can\$10 million (US\$6.7 million). The proposal is now being considered by the Canada Foundation for Innovation, a federal government agency, which is expected to make a decision by July.

If the foundation approves its Can\$3.8 million share, individual provinces are expected

to contribute another Can\$3 million. Diamond-mining companies should provide about Can\$1.5 million, with the remainder coming from several universities. “This is Canada’s biggest investment in equipment for Earth science,” says John Cassidy, a seismologist with the Geological Survey of Canada in Sidney, British Columbia.

US National Science Foundation (NSF) officials had hoped Mexico would fund a similar programme, in effect extending its USArray survey to the south as well as the north. Unfortunately, it seems that the money isn’t available. But some Mexican activity may still be possible, if scientists from the country put in compelling scientific proposals to use some of the seismometers currently operated for the NSF.

It's time to work together and stop duplicating conservation efforts ...

Sir — Myers *et al.*¹, in their new analysis of global biodiversity hotspots, recommend areas where conservation actions should be focused to minimize losses in the imminent extinction crisis. We strongly support initiatives to produce clear, efficient and practical goals for conservation to guide biodiversity planners and decision-makers in governments, agencies, conventions and non-governmental organizations (NGOs). However, as things stand there is only limited consensus on global conservation priorities at international level. We believe that the time is now right for scientists and practitioners to work together to develop a commonly adopted blueprint for action.

A key first step would be a structured debate to identify common goals; to pool data sets; and to agree on the contributions that non-equivalent measures of priority, such as diversity, endemism, threat, viability and ecological function should make. All this should lead to a powerful and cost-effective product. Most important, however, is that the sound analyses, breadth of expertise, and strength of consensus underpinning such an agenda would greatly increase the probability that decision-makers would adopt the resulting priority system, and take action to stem the loss of wild species and maintain critical ecosystem processes.

Currently there is much duplication of effort across organizations. Prioritization programmes with comparable goals to that of Myers and his collaborators at Conservation International are in progress at WWF-US², BirdLife International³, IUCN⁴, World Resources Institute⁵ and The Nature Conservancy⁶. All these are enormously costly, involving independent programmes to gather data, undertake analyses, and publish and advertise products. The result is considerable redundancy, and generation of competing rather than complementary priority sets (sets of areas that result from systematic prioritization). We appreciate that organizations have to develop specific products to meet their own objectives. Nevertheless, we believe that leading NGOs must cooperate in data-gathering, analysis and priority-setting if there is to be a widely supported strategy that the entire conservation community can work together to implement.

In our view, the analysis of Myers *et al.*¹ highlights several other issues that must be addressed in reaching consensus. First, a co-ordinated programme must take account of the many recent advances in systematic priority-setting⁷. The analysis by Myers *et al.* does not include new techniques allowing the identification of priority sets which, by

paying explicit attention to patterns of between-site complementarity, greatly boost the efficient representation of all mapped species and not just endemics targeted in pre-selected sites^{7–10}. These techniques also allow planners to justify the choice of particular areas, to determine the flexibility of site selection and to balance representation and conservation efficiency against a suite of socio-economic costs (such as land price, degree of land transformation, and human population density)^{7,9–11}.

Second, there remain several serious limitations to all current assessments of global priority. For instance, most conservation decisions take place at a finer geographical scale than that discussed by Myers *et al.*¹. Hence new techniques, which allow analyses to move freely between different scales, are needed before global analyses can be translated into effective action on the ground. At these finer scales there are already some encouraging collaborations among academic researchers and NGOs (see the letter by Fonseca *et al.*, below).

Equally pressing, priorities identified for one taxon may fail to reflect diversity in other groups¹² — hence analyses should be developed from independent data sets from several taxa, ideally including invertebrates: the taxon that represents most terrestrial biodiversity. Biological priorities must also be integrated with socio-economic considerations, but current developments in this area are still at a rudimentary stage.

Last, conservation practitioners and academic researchers alike are concerned that systematic prioritization focuses on patterns of species and community distribution, yet largely fails to address the conservation of key ecological and evolutionary processes which maintain those patterns¹³. There is growing evidence that conserving the pattern will not by itself guarantee the conservation of these processes. Yet we currently lack robust measures for quantifying the extent to which different areas contribute to core processes, or for evaluating the overall performance of priority sets in terms of process maintenance. In consequence, although it is claimed that the hotspots approach is cost-effective¹, problems concerning selection of areas, scale of analysis, socio-economic concerns and maintenance of key processes mean that areas of high conservation value may have been missed.

G. M. Mace

Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

A. Balmford University of Cambridge, UK

L. Boitani University of Rome, Italy

G. Cowlishaw Institute of Zoology, London, UK

A. P. Dobson Princeton University, USA

D. P. Faith Australian Museum, Sydney, Australia

K. J. Gaston University of Sheffield, UK

C. J. Humphries, R. I. Vane-Wright, P. H. Williams Natural History Museum, London, UK

J. H. Lawton Imperial College, Ascot, UK

C. R. Margules CSIRO Tropical Forest Research Centre, Australia

R. M. May University of Oxford, UK

A. O. Nicholls CSIRO Wildlife and Ecology, Australia

H. P. Possingham University of Adelaide, Australia

C. Rahbek University of Copenhagen, Denmark

A. S. van Jaarsveld University of Pretoria, South Africa

1. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. *Nature* **403**, 853–858 (2000).
2. Olson, D. M. & Dinerstein, E. *Conserv. Biol.* **12**, 502–515 (1998).
3. Bibby, C. J. in *Conservation Science and Action* (ed. Sutherland, W. J.) 176–201 (Blackwell Science, Oxford, 1998).
4. WWF & IUCN *Centres of Plant Diversity: A Guide and Strategy for their Conservation* 3 vols (IUCN Publications, Cambridge, 1994–97).
5. Ayensu, E. *et al. Science* **286**, 685–686 (1999).
6. The Nature Conservancy *Designing a Geography of Hope: Guidelines for Ecoregion-Based Conservation* (The Nature Conservancy, Arlington, 1997).
7. Margules, C. R. & Pressey, R. L. *Nature* **405**, 243–253 (2000).
8. Pressey, R. L., Humphries, C. J., Margules, C. R., Vane-Wright, R. I. & Williams, P. H. *Trends Ecol. Evol.* **8**, 124–128 (1993).
9. Faith, D. P. *et al. Forest Ecol. Mgmt* **35**, 251–260 (1996).
10. Williams, P. H. in *Conservation in a Changing World* (eds Mace, G. M., Balmford, A. & Ginsberg, J.) 211–249 (Cambridge Univ. Press, Cambridge, 1998).
11. Rodrigues, A. S. L., Tratt, R., Wheeler, B. D. & Gaston, K. J. *Proc. R. Soc. Lond. B* **266**, 1453–1460 (1999).
12. van Jaarsveld, A. S. *et al. Science* **279**, 2106–2108 (1998).
13. Balmford, A., Mace, G. M. & Ginsberg, J. in *Conservation in a Changing World* (eds Mace, G. M., Balmford, A. & Ginsberg, J.) 1–28 (Cambridge Univ. Press, Cambridge, 1998).

... following Africa's lead in setting priorities

Sir — It is vital to try to conserve biological diversity, but to do this we must know where it is. New databases mapping African biodiversity now offer this information^{1–3}.

Quantitative analyses⁴ of these data reveal geographical priorities for conservation broadly matching those generated at a global level by international conservation organizations^{5–7} (Fig. 1, overleaf). Although this approximate overlap with coarse-scale priorities is encouraging, we now need to move tropical priority-setting to finer scales that will enable conservation action. We suggest here three strategies for meeting this challenge: improving data; enhancing collaborations; and working closely with local decision-makers.

The main constraint on quantitative conservation priority-setting⁸ is that sufficient data are not available at the fine scale required for conservation implementation. The only long-term solution is to collect new biological data, compile and disseminate point locality information, and conduct basic taxonomy. Such work is chronically under-funded, yet is essential for conservation planning⁹. In the short term, the problem can be reduced by deductive modelling. Environmental data such as elevation, vegetation, rainfall and temperature have been modelled to fine resolutions — for example, 1-km grids — for the whole continent (see <http://edcftp.cr.usgs.gov>). If we compile information on habitat preferences, we can



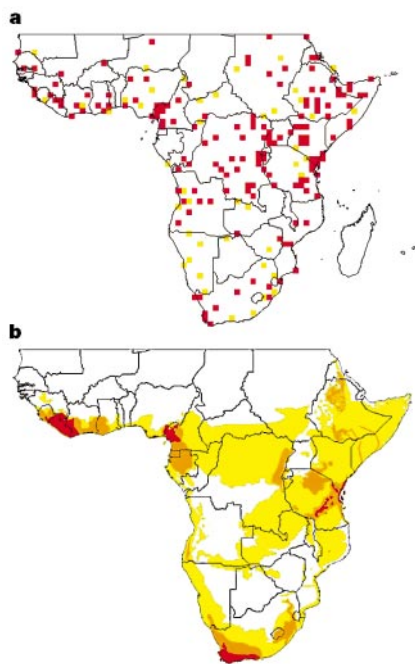


Figure 1 Conservation priorities for Sub-Saharan Africa. **a**, Quantitatively derived conservation priorities⁴ for ~4,000 species of bird, mammal, snake and amphibian, mapped on a 1° grid. Coloured cells depict the top 200 areas from which 97.5% of species mapped have been recorded. Red squares are irreplaceable because they contain the entire known distribution of one or more species; orange cells are flexible areas for which alternatives (not mapped) are available. **b**, Conservation International's Hotspots⁵; the World Wildlife Fund US's Global 200 most biologically important ecoregions⁶; and BirdLife International's Endemic Bird Areas⁷. Red, orange and yellow show areas of intersection between three, two and one system(s), respectively.

predict fine-scale species distributions by overlaying environmental data onto species-range maps, to identify areas where all of a species' habitat requirements are fulfilled³. Another short cut is to base conservation priorities on well-known taxonomic groups¹⁰. The problem with this is lack of knowledge of cross-taxon congruence¹¹ — for example, conserving birds may not be enough to protect biodiversity as a whole.

Attempts to address the lack of data on African biodiversity must go hand-in-hand with improved collaboration at all levels — see the letter from Mace *et al.* (page 393). Effective collaboration is urgently needed between the biological and social sciences, to incorporate human geography into the quantitative priority-setting process in the tropics¹². The development of parallel priority-setting initiatives is another symptom of the lack of effective coordination to date. Furthermore, difficulties have arisen over data dissemination and public access to information. Such tensions between data providers (for example, museums) and users (such as non-governmental organizations) can be eased by considering mutually beneficial collaborations. For instance, conservation groups could increase their funding

for the publication of biological data, and groups with mutual interests could collaborate on fund-raising to pay for data collection. Finally, much greater use should be made of existing collaborative networks¹³.

Effective translation of continental priorities into action depends fundamentally on consensus from local decision-makers. One way of forging this is through expert-based priority-setting workshops¹⁴ to assess key regional areas for conservation values in different taxonomic groups and to prioritize these areas across groups. From this synthesis, an integrated set of local priorities can be developed incorporating information on ecology, current and future threats, and landscape-level linkages. Essential components are a commitment to training, empowerment of local specialists, and repatriation of biodiversity information.

Such conservation-prioritization workshops have been held recently for the Upper Guinea region (December 1999) and for the Congo Basin (March 2000). The consensus forged by governmental, non-governmental and academic representatives from the countries in these regions has provided a solid base to translate priorities into action. We believe that our suggestions will considerably increase the chances of further progress while opportunities for effective conservation in Africa remain.

Gustavo A. B. da Fonseca

Conservation International Center for Applied Biodiversity Science, 2501 M Street NW, Washington DC 20037, USA, and Departamento de Zoologia, Universidade Federal de Minas Gerais, Avenida Antônio Carlos, 6627 Pampulha, Brazil

Andrew Balmford *University of Cambridge, UK*

Colin Bibby *BirdLife International, UK*

Luigi Boitani, Fabio Corsi *Istituto di Ecologia Applicata, Italy*

Thomas Brooks, Claude Gascon, Silvio Olivieri, Russell A.

Mittermeier *Conservation International Center for Applied Biodiversity Science, USA*

Neil Burgess *Wildlife Conservation Society of Tanzania*

Eric Dinerstein, David Olson *World Wildlife Fund, USA*

Lee Hannah *Conservation International, South Africa*

Jon Lovett *University of York, UK*

David Moyer *Wildlife Conservation Society, USA*

Carsten Rahbek *University of Copenhagen, Denmark*

Simon Stuart *International Union for the Conservation of Nature and Natural Resources, Switzerland*

Paul Williams *Natural History Museum, UK*

1. Burgess, N. *et al.* *Species* **30**, 16–17 (1998).

2. Lovett, J. C. *et al.* *Biodiv. Conserv.* **9**, 33–42 (2000).

3. Boitani, L. *et al.* *AMD African Mammals Databank: A Databank for the Conservation and Management of the African Mammals* (Istituto di Ecologia Applicata, Rome, 1999).

4. Williams, P. H. *WORLDMAP 4.1. Priority Areas for Biodiversity* (Natural History Museum, London, 1996).

5. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. *Nature* **403**, 853–858 (2000).

6. Olson, D. M. & Dinerstein, E. *Conserv. Biol.* **12**, 502–515 (1998).

7. Stattersfield, A. J. *et al.* *Endemic Bird Areas of the World: Priorities for Biodiversity Conservation* (BirdLife International, Cambridge, 1998).

8. Williams, P. H. in *Conservation in a Changing World* (eds Mace, G. M., Balmford, A. & Ginsberg, J.) 211–249 (Cambridge Univ. Press, Cambridge, 1998).

9. Balmford, A. & Gaston, K. J. *Nature* **398**, 204–205 (1998).

10. Bennun, L. A. & Fishpool, L. D. C. *Ostrich* (in the press).

11. Howard, P. C. *et al.* *Nature* **394**, 472–475 (1998).

12. Fjeldså, J. & Rahbek, C. in *Conservation in a Changing World* (eds Mace, G. M., Balmford, A. & Ginsberg, J.) 139–160 (Cambridge Univ. Press, Cambridge, 1998).

13. BCIS *Framework for Information Sharing* (Biodiversity Conservation Information System, Cambridge, 2000).

14. Olivieri, S. *et al.* *A Participatory Approach to Biodiversity Conservation: the Regional Priority Setting Workshop* (Conservation International, Washington DC, 1995).

Community groups could show Unesco the way

Sir — We approve of your recommendation that Unesco should focus more closely on its core goals (*Nature* **404**, 109; 2000). And we wish to go further with some concrete suggestions. Within the UN system, Unesco is unique in having a double representation. Each member State has a permanent diplomatic delegation representing its government, and, in addition, a national commission representing the academic and scientific community and 'civil society' — the grouping of all kinds of non-governmental organisations (NGOs) that exist to promote the interests of citizens.

The worldwide rise in influence of civil societies shows the way forward for Unesco. In all possible ways, Unesco should improve its relations with civil society organizations, and foster their development: closer cooperation with NGOs, universities and so on could actually lead to useful savings and increased efficiency.

Unesco should apply to its own functioning the principles of ethics and good governance. Its recurrent drifts into double standards (proclaiming lofty ideals while carrying out dubious internal practices) are no longer acceptable: Unesco should dedicate itself to promote a culture of evaluation.

Lastly, if we are serious about our social responsibilities, we scientists should see that science and technology remain firmly anchored inside Unesco.

Britain's return to Unesco in 1997 after leaving during the 1980s was a welcome move, but, for a proper implementation of the previous points, we do hope that the British National Commission for Unesco will be promptly reconstituted. From the French side, we eagerly look forward to fruitful discussions and stimulating debates with our British colleagues and partners. The pioneering contributions of Joseph Needham and Julian Huxley, at the birth of Unesco, have not been forgotten.

Our conviction is that the full restoration of this distinguished British tradition will be an important element in the renovation process of Unesco.

Gérard Toulouse

President of the Committee for Exact and Natural Sciences (French National Commission for Unesco), Laboratoire de Physique de l'ENS, 24 rue Lhomond, 75231 Paris, France

It won't come to the Crunch

Celebrate our new astronomical insight into the ultimate fate of the Universe.

The Runaway Universe: The Race to Find the Future of the Cosmos

by Donald Goldsmith

Perseus: 2000. 232 pp. \$25, £19.50

William H. Press

Like Lord Byron's deep and dark blue ocean, the cosmos rolls on. The cosmic expansion, as we see it relative to our own inconsequential position, proceeds at a speed of 70 kilometres per second (give or take) for every megaparsec of distance away from us. This number is the cosmologists' 'Hubble constant', its magnitude a subject of fierce dispute until the last decade, during which the likely range has so narrowed (65 to 75, say) as to make the remaining debate essentially scholastic — if no less interesting to specialists.

But is this year's Hubble constant the same as last year's? I don't mean this year's preferred measurement, but this year's actual value. Is the Universe expanding *more* or *less* rapidly this year than last year? And, if so, by how much? And can we even hope to measure what must surely be a tiny differential change?

Astonishingly, observational cosmologists are only now able to measure this effect — although there is still debate about how accurately. Even more astonishing is that the answer has turned the expectation of most theorists — not only cosmologists but also their cousins in particle and field theory — on its head.

The trick to doing the measurement is to increase the baseline: instead of comparing today's expansion with that of one year ago, compare it with that of 10^9 years ago by looking 'back in time' at very distant objects. But not just any objects will do; they must be objects with identical brethren in the here-and-now, because the differential measurement requires what amounts to a control population (astronomers refer to 'standard candles'). In just the past few years, two groups, one based at the University of California at Berkeley and the other at Harvard University, have perfected a technique for correcting the raw data from a certain kind of supernova, the 'Type Ia'. With this technique, they have created a superb control population for measuring the cosmic expansion and bringing within feasibility the measurement of its change with time.

Reputable theorists all expected expansion to be slowing down. That is because gravitation is attractive, so the initial expansion of the cosmos is slowed over time by the force of gravity fighting the expansion. Until recently, the debate was always framed thus:

is gravity going to win, leading to a 'Big Crunch', or is the expansion sufficiently robust that gravity will slow it only incrementally, leaving the cosmos free to 'coast' serenely to infinite expansion?

The Type Ia supernovae seem to show something much more peculiar. They do indeed indicate that the expansion is slowing, but not by as much as the known force of gravity should demand — for any plausible amount of matter. The evidence, found independently by the two groups, each with a lot of data, is that the known gravitation must be partly offset by a repulsive force (cosmologists delight in this much-repeated pun).

And so we look to the zoo of speculative repulsive forces, starting with the 'cosmological constant' invented by Einstein (and later repudiated by him), and moving to recent suggestions called 'quintessence' by their supporters (that is, a fifth essence, since physicists traditionally speak of four forces). By virtue of its seniority, although not its plausibility, the cosmological constant is given pride of place. Although now only a fractional offsetting of gravity on cosmological distances, it is destined (if it exists) to

become more and more dominant with time, driving the Universe to an exponential expansion that, in a mere 10^{11} years, will drive the galaxies so far apart from one another as to make them virtually isolated sub-universes.

Hence the "runaway Universe" in the title of Donald Goldsmith's excellent exposition of these and several related matters. Goldsmith is one of the great explainers of the astronomy of our time; and an explainer's craft is of a higher order than that of a mere popularizer. In this book, the science is in no way watered down or reduced to fuzzy analogies. Rather, the author explains, carefully and precisely, in ordinary language, the entire chain of discovery and logical reasoning that underlie the broad outline given above. This makes for some pretty dense prose from time to time, and some challenging graphs to study (no equations), but the text is enlivened, mercifully, by a requisite number of scurrilous personal stories, pieces of historical context and even a good measure of plain, dry wit.

About halfway through the book, when the implications of the new supernova data are becoming weighty, the author writes:



Dwarfed by space

The Helix nebula in the constellation of Aquarius is, at 450 light-years away, the closest planetary nebula to Earth. Shown above is its central star, a very hot white dwarf. Pictures of this and many other images obtained using the world's largest telescopes, including the Hubble Space Telescope, are to be found in *Majestic Universe: Views*

from Here to Infinity by Serge Brunier, translated by Storm Dunlop (Cambridge University Press, £25, \$30.95). The book charts the latest advances in cosmology, and the way in which the spatial and temporal limits of the cosmos have, in the past decade, been pushed back virtually as far as the Big Bang.

book reviews

"Let us pull ourselves from the slough of despond in which the implications of the new observations threaten to drown us. Rather let us lift our spirits by celebrating the new astronomical powers of insight that have brought us the latest news about the universe." A splendid idea. A few hours spent with this book is just the way to get those spirits hoisted. ■

William H. Press is at Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.

Stopping the tectonic conveyor belt

Dynamic Earth: Plates, Plumes and Mantle Convection

by Geoffrey F. Davies

Cambridge University Press: 1999. 470 pp.
£60, \$90 (hbk)/£24.95, \$39.95 (pbk)

David Bercovici

Geology students today are presented with the theory of plate tectonics as the grand model of how the Earth works. Centuries from now, however, it may well be viewed in the same way that we now view Kepler's laws of planetary orbits — as a set of descriptive, largely empirical rules that support a generalizable physical theory. In the case of Kepler's laws, the physical theory was Newtonian physics.

The physical theory that best accounts for plate tectonics is that of thermal convection in the mantle. In this theory, Earth's solid mantle, acting like a very viscous fluid over millions of years, loses heat as cold rock currents sink from near the surface to cool the mantle, while hot rising currents transport heat out of the mantle's deeper interior.

Convection is both elegant and universal, playing a part in everything from our cups of coffee to the interior machinations of stars. Mantle convection is thus a more fundamental theory of geology than is plate tectonics, although its presentation in most Earth-science texts is either muted or incorrect. In this regard, Geoffrey Davies' *Dynamic Earth* attempts to set the record straight: mantle convection is neither the oft-taught big wheel driving the tectonic conveyor belt nor merely small-scale circulation, unrelated to plates. As the book stresses repeatedly, the plates are mantle convection.

The book's underlying theme concerns the mantle's thermal boundary layers. These are relatively thin horizontal layers across which the high temperature of the mantle changes either to the colder temperatures of the atmosphere and oceans, at the upper boundary, or to the hotter temperature of the molten-iron outer core at the mantle's lower boundary. As these boundary layers retain most of the temperature contrasts of the



Magma event remembered

It is 20 years ago this month since Mount St Helens was transformed from a symmetrical cone of a mountain into a smoking crater. Its eruption killed 57 people and millions of animals, and decimated the landscape around it for hundreds of square miles. *Mount St. Helens: The Eruption and Recovery of a Volcano*

by Rob Carson (Sasquatch Books, \$19.95) tells the story of the event, from the preliminary, mile-high bursts of steam and ash seen throughout April, when a crater first began to open (inset, left), to the eruption itself (inset, right), its immediate aftermath, and the way in which life has slowly returned to the area.

system, they are also the primary sources of the thermal buoyancy that drives convection. The tectonic plates are thus formed from the upper thermal boundary layer, often sinking convectively along their edges in the form of cold subducting slabs which, in turn, pull the plates in their wake, hence driving plate motion. Hot mantle plumes, which cause areas of active volcanism such as occur at Hawaii and Iceland, probably emanate from the lower boundary layer.

In contrast, mid-ocean ridges, which occur in regions where plates are moving away from each other, are not pried apart by the hot upwelling currents of some giant, wheel-like convective circulation, but are pulled open passively, primarily from the force of sinking slabs. The author establishes the physical basis for this picture with clear explanations and simple equations, and uses it to explain the first-order observations, such as the surface features of the sea floor, seismic images, heat flow and mantle geochemistry.

Apart from various technical issues that I might quibble over, two philosophical problems lead me to recommend the book with caution. First, the author's presentation is

quite one-sided, drawing inordinately on his own two-dimensional view of mantle convection, and only mentioning alternative ideas if he can dispatch them quickly in a few sentences. Second, he gives the impression (and even explicitly states in the introduction) that our broad understanding of mantle convection will change little beyond the views presented in the book. In science there is no progress without obsolescence, and the suggestion that little significant work remains to be done only detracts from the excitement of an active scientific field. In the end, students and instructors alike should be wary of dogmatic books that claim to be the last word on a subject.

However, despite these cautionary statements, the book is an elegant and readable exposition with a clear message of how the plates and mantle probably work. The entrance of mantle-convection theory into mainstream Earth-science education is an important cause which this book champions admirably. ■

David Bercovici is in the Department of Geology and Geophysics, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA.

A Gouldian valediction, almost

The Lying Stones of Marrakech: Penultimate Reflections in Natural History

by Stephen Jay Gould
Jonathan Cape: 2000. 372 pp. £17.99, \$25.95

Henry Gee

Does the world really need another collection of Stephen Jay Gould's essays? *The Lying Stones of Marrakech* is the ninth, and its subtitle — *Penultimate Reflections in Natural History* — imply (nay, demand) a tenth. This will mark Gould's retirement (in 2001) from his 28-year stint as the unfailing monthly columnist for *Natural History* (the magazine of the American Museum of Natural History), even though *Lying Stones* contains miscellany such as sleeve notes to a CD of Mozart's music. (This gives Gould yet another opportunity to discuss contingency: what would have happened had Mozart lived to a ripe old age — what wonderful music would we have never heard? But what would the world have been like had he died even sooner?) But even Gould's barrel-scrappings are better than most other people's best efforts, so one mustn't grumble.

But when the *Ultimate Reflections* have coruscated into the empyrean welkin (I think I'm getting the hang of the polyglot agglutinative richness of the English language, don't

you?), will that be the end of Gould's contributions to literature? Contingency aside (any of us could be run over by a bus tomorrow), I can see a minor industry of Gouldian *recyclismus* looming. *Ultimate Reflections* will be followed by *Life's Wonderful Rich Grandeur*, a collection of the best of Gould's essays from his previously published cornucopia; further collections of essays they haven't squeezed into this collection; another collection of miscellany, and a collection of Gould's own introductions to these collections. By then, this small library will have appeared in paperback, giving the opportunity for another collection of this self-generating *oeuvre*. As T. S. Eliot never wrote: in the room the publishers come and go/ towards *absurdam, reductio*.

Amid this unending valediction, Gould will still produce the occasional straight book. There hasn't been an autobiography (*The Horologist in the Museum: Memoirs of a Mollusc-Hunting Man*), and a learned treatise on the obsessional statistics of baseball (*Say It Ain't So, Joe*) can't be far off. When Gould uses some extended anecdote about baseball as a prolegomenon (hah! I'm at it again!) to a homily on evolution, I put up with it. But in *Lying Stones* we have essays seemingly without number (well, only one, but it seems like more) on baseball, unalloyed. Now, I am a pronounced sportsphobe, and would have agreed with those patricians who damned such things as *panem et circenses* (kitchenware and circumcision, according to my own translation), fit only for those poor souls who must make do with sentimentality rather

than sentiment. And baseball is the basest of balls. Even duller than cricket, if that were possible, it has none of the spurious elegance of that dreary pastime: the *longueurs* of baseball consist of overweight men wearing silly uniforms, looking disagreeable, hugging their crotches and spitting a lot. It is a shame that Gould, a writer with such a deep intellectual life, feels the need to slum it in public. But *chacun à son gout*, as we Brits say. Again, mustn't grumble.

If I have one literary fault, it is digression, so I'll return to the rhetorical question I posed at the start and, in a spirit of paradox (because, hey, I'm that kinda guy) try to answer it. *Lying Stones* doesn't really say anything Gould hasn't said before, many times, and often more succinctly: the importance of contingency; the mysteries of the fossil record; man's inhumanity to man; celebrations of little-known (or misunderstood) figures from the pageant of science past, and so on. But each essay adds some tiny variation to the canon, and even if you know the ending, you can enjoy the minutely interesting detail along the way, not to mention Gould's commanding scholarship.

So, no, the world doesn't really need another collection of Gould's essays. But the world could also manage tolerably well without redbreasts whistling from garden crofts, or gathering swallows twittering in the skies: and (to be contingent) if Gould had become a tailor, or even a baseball player, and had never written a line, the world would be impoverished indeed. ■

Henry Gee is a Senior Editor at Nature.

The many faces of science

Defining Features: Scientific and Medical Portraits 1660–2000

by Ludmilla Jordanova
Reaktion Books & National Portrait Gallery:
2000. 192 pp. £14.95, \$24.95 (pbk)

Lisa Jardine

The cover of Ludmilla Jordanova's new book carries an engraving of the nineteenth-century astronomer William Herschel. The great man is portrayed in three-quarters profile, hair swept back from a high forehead (which for verisimilitude sports a small pimple), his large-collared greatcoat buttoned across his broad chest, sharp shadows angling his face. He is the archetypal figure of the eighteenth-century Romantic hero. Behind him a crescent moon rises above leafless trees in a dramatic night sky spangled with stars.

Appearances, however, as Jordanova explains, can be deceptive. This is not a portrait of a romantic dreamer, but of an



Covering all the bases: Joe DiMaggio, unwittingly a big hitter for evolutionary metaphor.

book reviews

energetic doer — a successful man of science. In the exhibition of portraits of scientists at the National Portrait Gallery in London (running until 17 September), which provided the occasion for Jordanova's book, the point is made even more clearly. Underneath the original Herschel print is a legend, written in swirling copperplate script. It tells us that the background shows, specifically, "part of the constellation of Gemini, with a telescopic aspect of the Georgium Sidus as it was discovered by Dr Herschel at Bath the 13th of March 1781 in consequence of which he was soon introduced to the most gracious patronage of His Majesty King George III".

The "Georgium Sidus" was the name Herschel gave — a compliment to his patron George III of England — to the new planet Uranus, which he discovered in the constellation Gemini using a seven-foot reflecting telescope that he had designed and built himself. This, then, is a tribute to the astronomer's greatest discovery, as well as a record of his features. And he is celebrated here as much for his mundane fund-raising success with the king — a keen amateur enthusiast for science — as for the spiritual high-mindedness of his astronomy.

Throughout history, people have been fascinated by what famous people look like. Before television, the portrait and the print satisfied this curiosity. Jordanova draws on the Wellcome Institute's extensive collection of Edward Jenner memorabilia to show how the face of the man who discovered inoculation against smallpox graced the walls and mantelpieces of even the humblest of a grateful public's homes. As in Herschel's case, what distinguishes the scientist from other celebrities in this respect is that, where the politician might hold a pen and the musician his instrument, the scientist's portrait is likely to allude to the scientific breakthrough itself. In Jenner's case, a cow and a milkmaid figure in almost every picture.

In what is probably the most interesting part of Jordanova's sometimes rather basic study, she discusses some of the more dubious ways in which women scientists have been portrayed for posterity. Here the tension between scientific achievement and the conventions of female virtue — passivity, docility and acquiescence, represented by demure dress and downcast eyes — is acute. Herschel's sister and scientific collaborator, Caroline, is represented demurely handing her brother a cup of tea, or primly bonneted, without a scientific instrument in sight.

Among the arresting portraits of Nobel prizewinner Dorothy Hodgkin that Jordanova reproduces, only Maggi Hambling's powerful 1985 oil painting shows the distinguished chemist and crystallographer with the tools of her trade. The others show her in film-star pose — her beauty apparently more significant than her scientific brilliance.

Even that most celebrated of women



Portrait of the scientist: Tom Phillips' video-based study of Royal Institution president Susan Greenfield.

ambassadors for the scientific profession (as current president of Britain's Royal Institution), the neuropharmacologist Susan Greenfield, is portrayed with emphasis on her good looks rather than her scientific expertise. Which makes the new portrait commissioned by the National Portrait Gallery, and forming the centrepiece of the current exhibition, all the more interesting.

For Greenfield's portrait, Royal Academy portraitist Tom Phillips, whose conventional portrait of mathematician Peter Goddard also features in the exhibition and book, has moved into new media. The portrait consists of a 15-minute loop, run on a Macintosh G4 computer, with a DVD driver. Its 22,500 frames are based on 169 drawings on paper, graphics onto screen, and short video sequences. The result is a compelling, elusive portrait that conveys physical traits and

mental innovativeness simultaneously. Here is a scientist whose gender is irrelevant, but whose intellectual curiosity is captured in the semi-abstract, constantly changing representation. Greenfield's shadowy face comes in and out of view through a filmy curtain of thought-provoking visual allusions.

The reproductions in this beautifully produced book capture much of the spirit of the exhibition it accompanies. Jordanova's text is occasionally ponderous, and she has lost some of the exhibition's buzz. But she succeeds in stimulating a fresh discussion of scientific portraiture. After this, we will all look with keener eyes at those familiar portraits that adorn the walls of the Royal Society or hang in splendour, tier upon tier, in the Royal College of Physicians. ■

Lisa Jardine is at Queen Mary and Westfield College, University of London, London E1 4NS, UK.

New Journals

This year, *Nature's* annual new journals review supplement will appear in the issue of 21 September. Publishers and learned societies are invited to submit journals for review, taking note of the following criteria:

- Journals must have first appeared during or after June 1998 and have published at least four separate numbers by the end of May 2000.
- Journals covering any aspect of science are eligible, although those dealing with clinical medicine and pure mathematics are excluded, as are newsletters and

publications of abstracts.

- Frequency of publication must be at least three times a year.
- The main language must be English.
- The deadline for submission is 5 June.

Please send at least four different issues (the first, the most recent and any two others) of each eligible title, or access details of any eligible electronic journal, together with full details of subscription rates, to: Isobel Flanagan, *Nature*, Porters South, Crinan Street, London N1 9XW, UK. Tel: +44 (0)20 7843 4542. e-mail: i.flanagan@nature.com

Many paths to enlightenment

Modern physics bears the imprint of Western and Asian philosophies.

Susantha Goonatilake

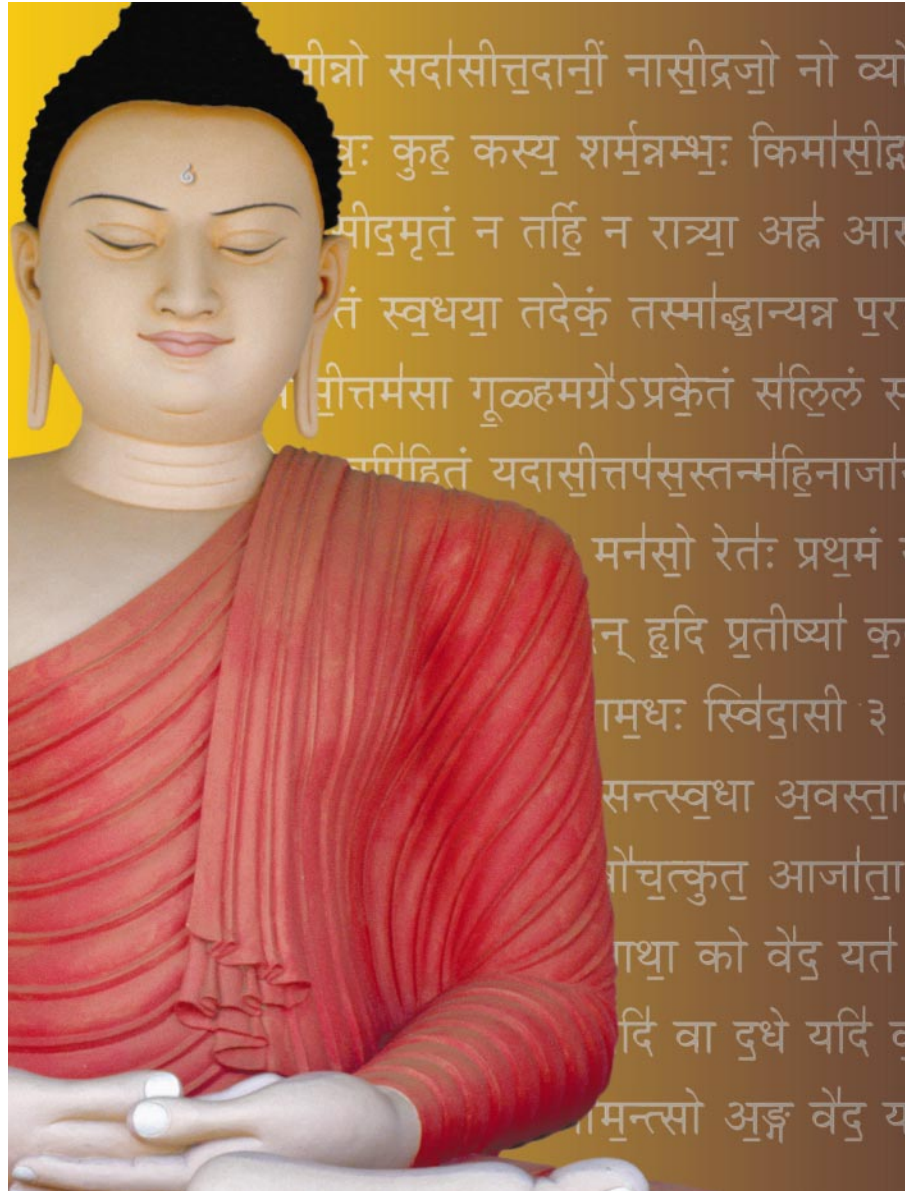
The discovery of the physics of relativity and quantum mechanics meant that the process of observation had to be radically reorientated. Relativity deemed all observational platforms to be equivalent, whereas in the quantum world, the act of observation itself changes the field of observation. The key theorists in the two realms, Albert Einstein and Erwin Schrödinger, were both deeply influenced by philosophy.

Einstein had been reading the philosophers David Hume and Ernst Mach just before he wrote his paper on special relativity. Their work helped to prepare the ground for Einstein's theories — he expressly credited Mach with influencing his thinking on special relativity, and wanted general relativity to conform to Mach's ideas. Einstein used Mach's principles to discount as a non-verifiable notion the possibility of two events happening simultaneously at different points in space.

In quantum physics — where common sense again breaks down — new philosophical insights are needed to transcend Newton's classical world. Schrödinger, whose equation defined the behaviour of particles and waves at the microscopic level, was influenced by Mach and Arthur Schopenhauer. In addition, Schrödinger declared himself a 'Vedantist', a category of Hindu philosopher.

Einstein and Schrödinger's influences resonated with south Asian philosophy. Both Hume and Mach believed that the self was a transient observational platform — Buddhists hold almost exactly the same view. Several commentators have noted the similarities between the thoughts of David Hume and Buddhism on the idea of the self. The philosopher Nolan Pliny Jacobson has observed that in both these philosophical viewpoints, which are separated by more than 2,000 years, "there is no thinker but the thoughts, no perceiver but the perceptions, no craver but the cravings. ... The similarity ... is striking". Jacobson's explanation for these parallels was the influence of Chinese culture on Europe during Hume's lifetime, which led to the spread of Buddhist ideas.

Mach, on the other hand, discovered Buddhism via Indian ideas. Some of his friends, such as Paul Carus and Theodor Beer, were Buddhists, and he contributed to Carus's journals *The Open Court* and *The Monist*. His first direct appreciation of a Buddhist position in relation to the relativity of the observer appears when he says: "But to ask that the observer should imagine himself as standing upon the Sun instead of upon the



Earth, is a mere trifle in comparison with the demand that he should consider the Ego to be nothing at all, and should resolve it into a transitory connection of changing elements." In 1913, he wrote of his happiness that Buddhism echoed his view of an unchanging ego as a "deception" and a "foolish notion".

As well as Mach, Schrödinger was influenced by Schopenhauer and Vedantism. Schopenhauer's debt to Buddhism is well known. And Walter Moore, Schrödinger's biographer, has noted that "the unity and continuity of Vedanta are reflected in the unity and continuity of wave mechanics". Ludik Bass, reviewing a biography of Schrödinger, noted that his "adherence went beyond

opinions, almost all that seems strange about him flowing from [Vedantist] philosophy". And Ramjit Nair, in a study of Schrödinger's philosophy, wrote that Vedantism was central to his thought — a strange combination of realism and absolute monistic idealism.

Einstein, in his obituary of Mach, emphasized the need for an interest in the theory of knowledge. This need is all the greater when a radical reorientation in thinking is required. In such instances, the teachings of other civilizations could provide a useful nudge to the scientific imagination. ■

Susantha Goonatilake is at the Center for the Study of Social Change, New School for Social Research, 5th Avenue, New York, New York 10011, USA.

Brochure

The apocalypse: a great day out for the whole family.

Joe Haldeman

Who among us is not enriched by a return to our roots? Billions of people and others have gone to great expense and trouble to find their way back — not just to the place where they were born, but to the various worlds to which they can trace their distant ancestors, affines and primary clones.

But not one in a million has gone all the way back.

For many centuries since Old Earth was condemned as officially uninhabitable, it has been sitting unused, a reliquary for archaeologists and historians, and a worst-case scenario for terraformers — notoriously, the most unmarketable piece of real estate in Sirius Sector. Now, at the dawn of the 31st 'Century', the Disney-Bertelsmann Consortium (DBC) is proud to announce that Terra is returning to usefulness, ready to take a new and important place in the human and humanized Universe.

Note: A 'century' was once a significant fraction of the human life span, and for the first 30 or so twyops after the discovery of consciousness, centuries were used as a primary chronological unit. Its

physio-symbolic origin had to do with the period of the planet around its primary and the number of manipulators primitive humans were born with. There were ten of them at the end of two 'arms'; and ten to the second power became a century.

Beginning at midnight, 1021.9445 Two-Pop, with allowances for relativistic contraction (31 December 3000 Old Style), DBC will be offering month-long Catastrophe Tours of Old Earth. A multi-gigacredit undertaking, Catastrophe Tours will thrill humans and others by recreating the unique sequence of disasters that finally resulted in the Diaspora and Rebirth.

Feel the horror!

As trillions of anaerobic microorganisms strangle in an atmosphere that slowly poisons itself with oxygen! DBC's unique *Time-Out™* field compresses eons into seconds, so that with your own eyes, or similar organs, you can watch the evolution of the tiny creatures that would become the temporary rulers of a transformed planet.

Watch the terrible glory!!

Of evolution in action, as multicellular creatures learn to eat each other while staking out more and more territory — from the

primeval depths up to the coral shallows, then further up onto land, growing from delicate amoebas to sponges, shrimp, fish — and then from clever little lizards finally into huge lumbering reptilian monsters, and then ...

See the terror!!!

As a mighty asteroid crashes into the Yucatan Peninsula! With DBC's *TimeOut™* field working in reverse, follow the shockwave in excruciating slow motion, as it circles the planet, destroying thousands of species and paving the way for the temporary primacy of mammals! For species with sufficiently heavy armour, DBC offers the option of 'surfing' the shockwave — watching in real time as it pulverizes everything in its path. (Participants must be able to withstand impulses greater than 1000 standard joves and temperatures greater than molten iron.)

Or, for somewhen a bit cooler...

Freeze off your extremities!!!!

As ice fields expand from the Poles nearly to the Equator! Again, DBC's *TimeOut™* field transforms what was once a glacially slow strangulation into a breathtaking crystalline transformation. An optional winter sports package is available for suitably hardy species, although DBC takes no responsibility for your safety. Some of the animals are big, and all of them are hungry.

Thrill to the 'Four Horsemen'!!!!!!

As the 'human' race, in its primal incarnation, is decimated time and again by war, pestilence, famine and plague. They aren't real, of course, by modern standards, but their suffering is real enough!

Laugh at the 'Final Trump'!!!!!!

Whose humorous double meaning will be explained, when a combination of ultimate weapons with greed finally ends war and suffering — by eliminating nearly all potential victims! Customers with a sufficiently developed appreciation of irony, and lovers of symmetry, will be amused by the rightness of the ultimate disappearance of oxygen, and the return of the planet to its original anaerobic owners.

Disclaimer: No advanced life forms are harmed by this entertainment. The only creatures that survive on Earth are a few hardy, rapidly mutating species that thrive on disaster.

Joe Haldeman has written more than 20 novels, and four collections of short stories and poetry. His latest novel, *Forever Free* (a sequel to *The Forever War*), is published by Orion in the United Kingdom and Ace in the United States.

"WATCH THE TERRIBLE GLORY"

**SEE THE TERROR
FEEL THE HORROR**

CATASTROPHE TOURS

LAUGH AT THE FINAL TRUMP !!!

Disclaimer: No advanced life forms are harmed by this entertainment. The only creatures that still survive on Earth are a few hardy, rapidly mutating species that thrive on disaster.

1021.9445 Two-Pop

JACEY

Loophole for snowball Earth

Bruce Runnegar

The snowball Earth hypothesis posits an ice-covered planet. New climate simulations of 'snowball' conditions allow ice-free equatorial oceans that may be crucial for a theory about early animal evolution.

Snowball Earth is a script for global catastrophe that rivals giant-impact theories in the severity of its postulated environmental effects. In the 'hard' version of this hypothesis¹, low concentrations of atmospheric carbon dioxide (around 150 μ bar) and fainter sunlight (by some 6%) allowed polar ice caps to grow until reflected sunlight cooled the Earth and the oceans froze to a depth of about one kilometre. According to proponents, this happened at least twice during the Cryogenian period of the late Precambrian (850–590 million years ago) and perhaps more frequently a billion-and-a-half years before that^{2,3}. Recovery from these episodes depended on CO₂ emissions from volcanoes, and consequent greenhouse warming.

The effect on the biosphere of snowball events is thought to have been catastrophic. Carbonates immediately below and above the glacial deposits record carbon-isotope ratios characteristic of Earth's mantle^{1,4}, rather than of life processes, implying that oceanic photosynthesis was effectively eliminated during each snowball event. The result of this and anoxic conditions beneath the ice should have annihilated most kinds of eukaryotic life — that is, almost everything except bacteria. So successive snowball Earths should represent bottlenecks in the evolutionary history of eukaryotes through which comparatively few organisms passed. Conversely, the final disappearance of snowball conditions may have permitted and even stimulated the Cambrian explosion of complex multicellular life that began some 565 million years ago.

On page 425 of this issue, Hyde *et al.*⁵ challenge the 'hard' snowball Earth hypothesis. Their energy-balance and general-circulation climate models allow open water in equatorial regions to coexist with snowball Earth conditions elsewhere. According to their calculations, massive continental ice sheets up to 5 km thick flowed outwards into low latitudes. However, providing that atmospheric CO₂ was above a plausible level (roughly double the present amount), these flows did not create global oceanic ice shelves. Instead, they calved and melted to leave a narrow circum-equatorial refugium that might have included some access to coastal zones.

This 'softer' picture of a snowball Earth is similar to the original version proposed by

Kirschvink a decade ago⁶, and it allows for the evolution of the biosphere to proceed in a more orderly fashion than does the 'hard' alternative. But which of those views is likely to be correct?

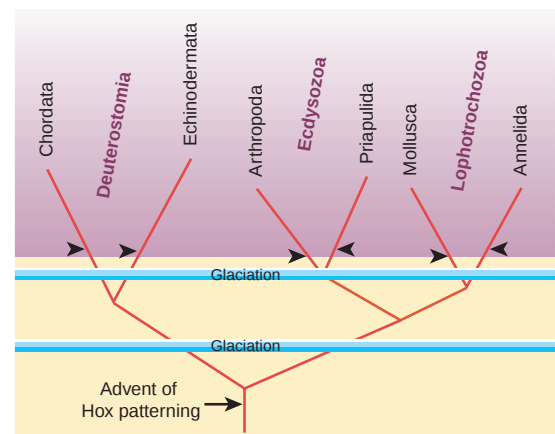
A key question not addressed by Hyde *et al.* is whether enough CO₂ could accumulate in an atmosphere in contact with the global ocean to end a snowball Earth event. Hoffman *et al.*¹ assumed that more than a tenth of a bar of CO₂ would be required to overcome the increased reflectivity of an ice-covered Earth, so it is not a trivial matter.

Presumably, all continental silicates and shallow-water carbonates were protected from chemical weathering by very cold ice cover, so the usual controls on CO₂ build-up would not have applied⁷. If total alkalinity remained relatively constant, only a fraction of the carbon that needed to accumulate in the atmosphere could have been absorbed by the ocean. However, such an ocean would have been acidic, freezing, undersaturated in calcium carbonate, and inimical to the formation of mineral skeletons. This is one possible reason why the

Box 1 How to make an animal — indirectly

In a provocative series of articles, summarized in ref. 9, Peterson, Cameron and Davidson have proposed a new model for the origins of the disparate body plans of the various bilaterian metazoan phyla — primitively mobile animals that have an anterior–posterior axis, a central nervous system, and tissues and organs derived from a middle body layer.

Molecular phylogenies segregate all bilaterian phyla into three large groups: deuterostomes, ecdysozoans and lophotrochozoans, shown in the figure here with some distinctive component phyla. The last two are sister groups and they jointly share the last common ancestor of all living bilaterians with the deuterostomes. That animal, argue Peterson *et al.*, had a free-living larval stage that bore little or no resemblance to the adult. The adult body was built from undifferentiated 'set-aside' cells that were free of growth constraints imposed on all other parts of the embryo. These cells could thus participate in a more sophisticated developmental process, involving Hox gene expression patterns, that is



characteristic of all bilaterian phyla.

The key point here is that similar bilaterian larvae are thought to have given rise to very different adult body plans. A classic case is the dissimilarity of adult molluscs and sipunculan worms and the great similarity of their larvae. But here lies the difficulty. Molluscs and sipunculans are lophotrochozoans, so their origins lie well above the common bilaterian ancestor in the metazoan tree. Somehow, deployment and patterning of pre-existing set-aside cells had to be delayed until the various lineages leading to the bilaterian phyla had been

separated for long periods of time. For example, distantly related members of the three bilaterian groups appear abruptly in the fossil record at the beginning of the Cambrian around 545 million years ago (arrows). However, molecular clocks¹⁰ suggest that the common ancestor of all three groups lived before the Cryogenian glaciations. Early-diverging bilaterians may therefore have been kept in 'larval mode' after the invention of set-aside cells and Hox cluster patterning by being forced to survive one or more snowball Earth glaciations in open-ocean environments.

B.R.

Cambrian explosion might have been postponed (Box 1).

But there is another outcome of the work of Hyde *et al.*⁵. Let us accept, tentatively, that there were two global glaciations that occurred about 200 and 50 million years before the beginning of the Cambrian⁴. If so, each could have had one of three effects on the biosphere, in decreasing order of severity.

A snowball bottleneck. Conditions: global oceanic ice cover for around ten million years; anoxic oceans; almost complete loss of photosynthesizing biomass; extinction of all but a few lineages of eukaryotes, which survived mainly in geothermal oases on the continents. Probable consequences: a massive bottleneck in eukaryotic lineages; survivors that were predominantly terrestrial in origin.

A blue-water (pelagic) refugium, such as emerges from the paper of Hyde *et al.* Conditions: massive continental ice sheets and marginal ice shelves, but an open-water circum-tropical ocean with unglaciated volcanic islands; almost complete destruction of terrestrial biota and shallow-water, bottom-dwelling life; oxygenated oceans; prolific pelagic biomass in the nutrient-rich equatorial waters. Probable consequences: snowball episodes filtered eukaryotic lineages by allowing the survival of organisms that inhabited either the pelagic, open-ocean realm, or ocean-floor hydrothermal vents.

A uniformitarian outcome. Conditions: global refrigeration, but minimal impact on the biosphere because habitat loss was offset by positive effects such as the increased solubility of oxygen in sea water. Probable consequences: little or no evidence for snowball episodes in the evolutionary histories of post-Cryogenian marine eukaryotes.

The possibility that early metazoan (multicellular animal) lineages went through a pelagic filter may help to overcome a fundamental difficulty with one explanation for the origin of metazoan body plans^{8,9}. According to this view, described in more detail in Box 1, a unique innovation — the evolutionary origin of larval 'set-aside' cells — provided the canvas on which a wide range of animal body plans could ultimately be sketched. But to generate the main metazoan groups by this method it is necessary to separate lineage-splitting in time from body-plan specification. In other words, some environmental filter was required to maintain early metazoans in 'larval mode' after they had invented set-aside cells. This enabled them to diversify into well-separated lineages that ultimately became the independent sources of radically different body plans. A blue-water refugium could have served this purpose. If so, the Cambrian explosion might have begun with pelagic organisms discovering or rediscovering the sea floor.

Bruce Runnegar is in the IGPP Center for Astrobiology and the Department of Earth and

Space Sciences, University of California, Los Angeles, California 90095-1567, USA.
e-mail: runnegar@ucla.edu

1. Hoffman, P. F. *et al.* *Science* **281**, 1342–1346 (1998).
2. Evans, D. A. *et al.* *Nature* **386**, 262–266 (1997).
3. Schmidt, P. W. & Williams, G. E. *Earth Planet. Sci. Lett.* **172**, 273–285 (1999).
4. Kennedy, M. J. *et al.* *Geology* **26**, 1059–1063 (1998).

5. Hyde, W. T., Crowley, T. J., Baum, S. K. & Peltier, W. R. *Nature* **405**, 425–429 (2000).
6. Kirschvink, J. L. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 51–52 (Cambridge Univ. Press, 1992).
7. Berner, R. A. & Cladeira, K. *Geology* **25**, 955–956 (1997).
8. Davidson, E. H. *et al.* *Science* **270**, 1319–1325 (1995).
9. Peterson, K. J. & Davidson, E. H. *Proc. Natl Acad. Sci. USA* **97**, 4430–4433 (2000).
10. Wang, D. Y.-C. *et al.* *Proc. R. Soc. Lond. B* **266**, 163–171 (1999).

Cancer

New link in a web of human genes

Jean Y. J. Wang

Biologists have solved complex problems through genetics for more than a century. The simple approach of cataloguing genetic mutants often reveals interesting and important relationships between genes. By applying this principle, several groups^{1–4}, including two whose reports appear on pages 473 and 477 of this issue^{2,3}, have now established a functional link between *ATM* and *NBS1* — two genes involved in human diseases.

ATM is mutated in the disease ataxia

telangiectasia (AT), the symptoms of which include degeneration of a particular brain region (the cerebellum), immune dysfunction, sterility, sensitivity to radiation, and increased risk of cancer. *NBS1* is mutated in a rare disease called Nijmegen breakage syndrome (NBS). Patients with NBS have symptoms similar to those seen in AT. Although AT and NBS can be distinguished by clinical criteria, cells of AT and NBS patients show striking similarities, including the increased breakage of chromosomes.

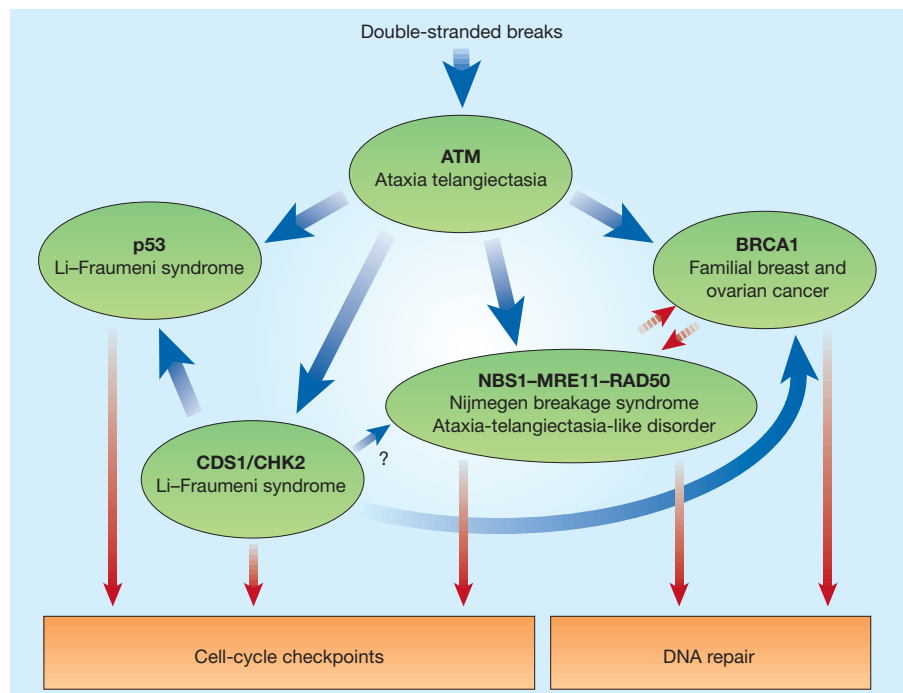


Figure 1 A network of proteins, implicated in human cancer, that regulate cellular responses to DNA damage. Double-stranded breaks (DSBs) in chromosomal DNA can activate a network of phosphorylation events to regulate DNA repair and progression through the cell-division cycle. Phosphorylation events are indicated by blue arrows. The protein kinase ATM is activated by DSBs to phosphorylate p53 (refs 1, 5, 6), CDS1/CHK2 (ref. 7), NBS1 (refs 1–4) and BRCA1 (ref. 9). (Mutations in the genes encoding these proteins are characteristic of the diseases indicated.) Further phosphorylation of p53 (refs 13–15) and BRCA1 (ref. 16) by CDS1/CHK2 might reinforce the actions of ATM. It will be interesting to know whether CDS1/CHK2 might also phosphorylate NBS1. The NBS1-MRE11-RAD50 complex also interacts with BRCA1, perhaps to coordinate DNA repair (dashed red arrows). Phosphorylation of p53, CDS1 and NBS1 activates cell-cycle checkpoints, which prevent the duplication and propagation of damaged DNA. Phosphorylation of NBS1 and BRCA1 may regulate the repair of DSBs. Together, the genes in this network protect the integrity of the genome, and mutations in the network predispose people to cancer.

A chromosome is broken when a double-stranded break (DSB) occurs in its DNA. The formation of DSBs is quite usual in many circumstances; for example, it is a necessary step in the production of gametes (sperm or eggs) and functional lymphocytes (immune cells). In normal gametes, lymphocytes or other cells, DSBs are processed and repaired. AT and NBS patients, however, have a general defect in dealing with DSBs that might account for their symptoms.

The genes mutated in AT and NBS — *ATM* and *NBS1*, respectively — are both involved in the cellular response to DSBs. The wild-type *ATM* protein is a protein kinase, an enzyme that adds phosphate groups to (phosphorylates) other proteins. It is stimulated by DSBs to activate 'check-points' — mechanisms by which a cell prevents itself from copying or passing on damaged DNA⁵⁻⁷. Wild-type *NBS1*, by contrast, is one component of a trimeric protein complex (the other subunits being *MRE11* and *RAD50*) that is recruited to sites of DSBs to process the broken DNA ready for repair⁸.

The similarities between AT and NBS led the authors of the four new papers¹⁻⁴ to investigate whether there might be any functional links between the proteins affected — and indeed there are. The four groups show that DSBs stimulate *ATM* to phosphorylate *NBS1* at several amino acids, and that these phosphorylation events are required for the proper cellular response to DSBs. These results are important for several reasons. First, they show that the phosphorylation of *NBS1* by *ATM* is essential in inhibiting DNA synthesis upon DSB formation. Second, they establish that *ATM* has a role not only in preventing cells from propagating damaged DNA, but also — through *NBS1* — in the processing and repair of DSBs. Finally, the biochemical relationship between *ATM* and *NBS1* nicely explains the similarity between AT and NBS.

The phosphorylation of *NBS1* is complicated. Lim *et al.*¹ and Gatei *et al.*⁴ show that, after treatment of cells with ionizing radiation, the serine residue at amino-acid position 343 of the protein is phosphorylated, whereas Zhao *et al.*² find both serines 278 and 343 to be phosphorylated after irradiation. Wu *et al.*³, by contrast, observe that serines 343, 397, 432, 509 and 604 in *NBS1* are phosphorylated even before irradiation; after irradiation, the amount of phosphorylation is increased at these sites. *ATM* itself phosphorylates *NBS1* at at least four sites¹⁻⁴: serines 278, 343, 397 and 615. The authors investigated the importance of these sites by mutating each of them, individually, to alanine to abolish phosphorylation. Irradiated cells carrying the serine 343/alanine *NBS1* mutant were defective in inhibiting DNA synthesis^{1,2}. All four of the mutants showed some, but diminished, ability to protect NBS cells from radiation^{2,3}. So, mutation at any

one of these four serines (leading to a failure to phosphorylate these positions) compromises the biological function of *NBS1*.

Two models could explain the requirement for phosphorylation of several residues in *NBS1*. The first invokes sequential events: the essential phosphorylation occurs only after earlier phosphorylation events. In the second model, *NBS1* has several different functions, each of which is regulated by a distinct phosphorylation event. More data will be needed before we know which, if either, of these two models is correct.

ATM can phosphorylate several proteins that are mutated in human cancer^{1-7,9} (Fig. 1). Interestingly, however, among the known targets of *ATM*, only *NBS1* (and its partner *MRE11*) are involved in diseases that are similar to AT. *MRE11* is mutated in a rare form of AT-like-disorder (ATLD)¹⁰. On the basis of genetic information and the enzyme-substrate (*ATM*-*NBS1*) relationship described in the new papers, one might conclude that *NBS1* is the essential target of *ATM*. But if this were true, *ATM* would have no function without *NBS1*. Not so — although *ATM* cannot inhibit DNA synthesis without *NBS1* (refs 1, 2) and *MRE11* (ref. 10), it can phosphorylate and activate the tumour-suppressor protein *p53* in the absence of these proteins^{1,10}. Under some conditions, *NBS1* can also function without *ATM*: Wu *et al.*³ detected increased phosphorylation of *NBS1* in cells from AT patients (in which *ATM* is mutated) on exposure of the cells to ultraviolet radiation. But, although the mutually dependent relationship between *ATM* and *NBS1* does not apply to all biological situations, disruption of this link is likely to be the root cause of the defects in AT, NBS and ATLD.

The biochemical connection between *ATM* and *NBS1* also establishes a new link in a network of genes that are implicated in human cancer. This network includes the genes encoding *NBS1*, *MRE11*, *BRCA1*, *p53* and *CDS1/CHK2* (ref. 11; Fig. 1). These proteins are joined in a web of phosphorylation events that are triggered by *ATM*. This web

has at least two biological functions — the repair of DSBs and the regulation of cell proliferation — and therefore has the important role of safeguarding the genome. Consequently, defects in this network can drive cancer development.

Although the primary function of this network is to repair DSBs and so promote cell survival, it also activates apoptosis, a cell suicide programme, which eliminates damaged cells and protects the organism. For example, *ATM* can activate apoptosis in neurons that suffer from DSBs¹². Yet cells should be reluctant to die upon DSB formation, as DSBs occur so easily during DNA replication and during the formation of gametes and lymphocytes. A DSB-activated network must be able to assess the damage and discriminate between continued proliferation versus death. The known links in this network of cancer genes do not reveal the basis for such discrimination, so yet more proteins must be involved. But the more we know about this complicated web, the better are our chances both of understanding the underlying causes of cancer and of developing effective therapies. ■

Jean Y. J. Wang is in the Department of Biology and the Cancer Center, University of California, San Diego, La Jolla, California 92093-0322, USA. e-mail: jywang@ucsd.edu

1. Lim, D.-S. *et al.* *Nature* **404**, 613–617 (2000).
2. Zhao, S. *et al.* *Nature* **404**, 473–477 (2000).
3. Wu, X. *et al.* *Nature* **404**, 477–482 (2000).
4. Gatei, M. *et al.* *Nature Genet.* **25**, 115–119 (2000).
5. Banin, S. *et al.* *Science* **281**, 1674–1677 (1998).
6. Canman, C. E. *et al.* *Science* **281**, 1677–1679 (1998).
7. Matsuoka, S., Huang, M. X. & Elledge, S. J. *Science* **282**, 1893–1897 (1998).
8. Paull, T. T. & Gellert, M. *Genes Dev.* **13**, 1276–1288 (1999).
9. Cortez, D., Wang, Y., Qin, J. & Elledge, S. J. *Science* **286**, 1162–1166 (1999).
10. Stewart, G. S. *et al.* *Cell* **99**, 577–587 (1999).
11. Bell, D. W. *et al.* *Science* **286**, 2528–2531 (1999).
12. Chong, M. J. *et al.* *Proc. Natl Acad. Sci. USA* **97**, 889–894 (2000).
13. Hiraio, A. *et al.* *Science* **287**, 1824–1827 (2000).
14. Chehab, N. H., Malikzay, A., Appel, M. & Halazonetis, T. D. *Genes Dev.* **14**, 278–288 (2000).
15. Shieh, S.-Y., Ahn, H., Tamai, K., Taya, Y. & Prives, C. *Genes Dev.* **14**, 289–300 (2000).
16. Lee, J.-S., Collins, K. M., Brown, A. L., Lee, C.-H. & Chung, J. H. *Nature* **404**, 201–204 (2000).

Star formation

Three's a crowd

Alan P. Boss

Looking at the night sky with the naked eye gives the impression that nearly all stars are single stars, much like our Sun. In reality, roughly half of these stars are binaries — two stars revolving around a common centre of mass. Studies of star formation have focused largely on single stars. In fact, the theory of how single stars form is so well developed and accepted that a 'standard

model' for their formation has thrived for well over a decade¹. However, startling new results presented at a meeting in Potsdam last month* suggest that the star-formation process may only rarely lead directly to the formation of single protostars — instead, protostars and young stars appear to form

*The Formation of Binary Stars International Astronomical Union Symposium 200. Potsdam, Germany, 10–15 April 2000.

primarily in groups of two or more, the latter of which could then disintegrate to produce single stars.

Because it is generally difficult to detect binary companions, their estimated frequency has increased with time as observational techniques improved. Nearby stars are the easiest to examine, and studies show that about half of all the primary stars with masses ranging from that of the Sun to ten times smaller (G to M dwarfs; see Fig. 1) have binary companions (M. Mayor, Geneva Obs.). Close binary companions are also turning out to be common for the very-low-mass brown dwarf stars (G. Basri, UC Berkeley). Even the much older stars in the galactic halo have the same binary frequency as the nearby disk stars (D. Latham, Harvard-Smithsonian Center Astrophysics.).

However, there is evidence that the more massive OB stars have more than twice as many companions on average as solar-mass stars, and even this estimate is likely to be a lower bound because of incompleteness (T. Preibisch, MPI Radioastron., Bonn). The rate of multiple systems is also high for the more massive stars.

The binary frequency for young stars is often considerably higher than that of the much older stars in the solar neighbourhood². Because young stars are more distant and harder to scrutinize, their binary fraction can only increase as techniques improve. In addition, the spectral variability of young stars makes them poor targets for spectroscopic searches for binary companions; even so, spectroscopic binaries seem to be at least as common among young stars as around older stars.

Herbig Ae-Be stars are young stars somewhat more massive than the Sun, and they

Main sequence stars	
Spectral type	Approximate mass (solar units)
O	60–16
B	16–3
A	3–1.5
F	1.5–1
G	1–0.8
K	0.8–0.5
M	0.5–0.08
Pre-main sequence stars	
Designation	Approximate mass (solar units)
Herbig Ae-Be	6–2
T Tauri	2–0.2

Figure 1 Star rating. In main sequence stars, the source of energy is the fusion of hydrogen into helium. Most stars are main sequence. The core of a pre-main sequence star is hotter and denser than that of a protostar, but not enough for hydrogen fusion to dominate. Solar units are multiples of the mass of the Sun. Spectral type is an indication of the surface temperature of the star, with O the hottest type at >25,000 K, and M the coolest at <3,500 K.

appear to have at least twice as many binary companions as older stars (J. Bouvier, Grenoble Obs.). T Tauri stars are young stars with solar masses, and in many star-forming regions they have twice as many binary companions as the nearby stars³. Amazingly, T Tauri itself, the prototypical single young star, has turned out to be not only a binary, but a member of a triple system.

Observations of the large-scale molecular gas outflows and smaller-scale optically visible jets that accompany young stars lead to the remarkable conclusion that essentially all such outflows originate in binary or multiple star systems (B. Reipurth, Univ. Colorado, Boulder). In fact, the prototypical outflow source L1551 IRS5, originally thought to be a single protostar, is now known to be not just a binary protostar⁴, but a member of a triple system, like T Tauri.

Furthermore, multiplicity emerges very early in the star-formation process — the very youngest protostars are often members of binary, triple and even quadruple systems (Fig. 2), and single protostars seem to be rare in star-forming regions (L. Mundy, Univ. Maryland, College Park). Some of these multiple systems consist of roughly equidistant protostars, a highly unstable configuration whose evolution must involve the ejection of single stars, leaving behind binaries and hierarchical multiple systems. In others, loss of gas or interactions with nearby protostars could lead to a gravitationally unbound system and hence to disruption.

Observers have a habit of disproving otherwise perfectly good theories, leading some theorists to avoid their company, but the Potsdam meeting proved to be a love-in. Just as observers found themselves with an abundance of binary and multiple systems of young stars, so theorists found themselves more or less in agreement that fragmentation, a mechanism capable of producing copious binary and multiple systems, seems to be the only game left in town. Fission of a rapidly rotating star leads to the equatorial ejection of matter rather than to the formation of two separate stars (R. Durisen, Indiana Univ.), and infrequent capture of young protostars through gravitational interactions between the protostars and their circumstellar disks cannot explain the circumbinary disks frequently found around young binary stars (C. Clarke, Cambridge Inst. Astron.).

Fragmentation, on the other hand, seems to meet many observational criteria: break-up during the collapse phase prior to protostar formation readily explains the early appearance of multiplicity, coequality of binary pairs, a wide range of separations, eccentric orbits, and circumbinary disks. High-spatial-resolution models of collapsing, fragmenting clouds

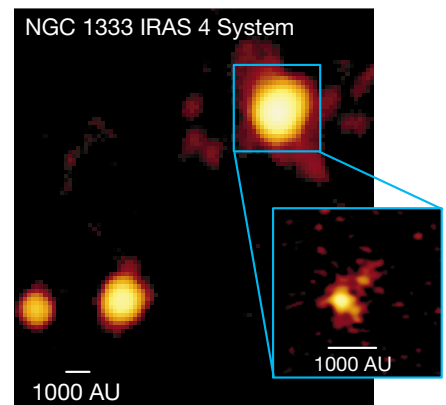


Figure 2 Millimetre-wave image of a cluster of four young protostars in a nearby star-forming region, NGC1333. Inset, enlargement of overlapping stars. The image depicts dust continuum emission, obtained with the Berkeley-Illinois-Maryland Array (BIMA) at a wavelength of 2.7 mm. (Courtesy L. Mundy, Univ. Maryland, College Park.)

suggest that the inclusion of previously neglected physical effects, such as turbulence (A. Burkert, MPI Astron., Heidelberg; R. Klein, UC Berkeley) and magnetic fields (A. Boss, Carnegie Inst.), seem to enhance the formation of binary and multiple protostars, the latter of which could then decay to produce single stars (R. Larson, Yale Univ.). Rapid dissipation of magnetic support through processes such as ambipolar diffusion may be a key ingredient for binary star formation (F. Shu, UC Berkeley).

Fragmentation thus seems to have been crowned as the dominant mechanism for binary star formation, as appropriate for a meeting held in the imperial Prussian city of Potsdam. However, there are implications for the further discovery of extrasolar planetary systems. If most single stars have to experience a close encounter with another protostar in the process of being ejected from a multiple system, their circumstellar disks — or at least their outer regions — will be in jeopardy from tidal forces during the encounter⁵. It might then turn out that wide binaries, which did not result from the decay of an unstable multiple system, could be more hospitable to the formation of planetary systems than ejected single stars.

Alan P. Boss is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington, DC 20015-1305, USA.

e-mail: boss@dtm.ciw.edu

- Shu, F. H., Adams, F. C. & Lizano, S. *Annu. Rev. Astron. Astrophys.* **25**, 23–72 (1987).
- Mathieu, R. D. *Annu. Rev. Astron. Astrophys.* **32**, 465–530 (1994).
- Ghez, A. M., McCarthy, D. W., Patience, J. L. & Beck, T. L. *Astrophys. J.* **481**, 378–385 (1997).
- Rodriguez, L. F. *et al. Nature* **395**, 355–356 (1998).
- Clarke, C. J. & Pringle, J. E. *Mon. Not. R. Astron. Soc.* **249**, 584–587 (1991).

Genomic imprinting

Silence across the border

Wolf Reik and Adele Murrell

Imprinted genes in mammals are expressed from only one chromosome: either the gene inherited from the mother or that from the father is silenced. Such genes include *Igf2* and other genes involved in growth control; here, imprinting might work to balance maternal and paternal demands on the size of an embryo. We know of several mechanisms by which imprinting can be achieved. And now the work of several groups^{1–4}, including two whose reports appear on pages 482 and 486 of this issue^{1,2}, allows us to add yet another weapon — chromatin boundaries — to the imprinting arsenal.

The differential labelling of cytosine–guanine base pairs with methyl groups distinguishes the two copies of an imprinted gene. This differential methylation usually originates in the parental germ cells — in egg or sperm⁵. It is generally assumed that methylation of a gene ensures that it cannot be transcribed. Indeed, in mice deficient in methylation, some imprinted genes are expressed from both parental chromosomes⁶. Surprisingly, however, other genes then become silent on both chromosomes⁶. Clearly, we do not know all there is to know about the phenomenon of imprinting.

The four groups^{1–4} studied the *H19* and *Igf2* genes, which are closely linked imprinted genes (Fig. 1) that are part of an imprinted gene cluster. They are oppositely imprinted: on the maternal chromosome, *H19* is turned on and *Igf2* is turned off; the opposite is true for the paternal chromosome. Both genes share a set of enhancers, or regions that control gene expression, which are located downstream of *H19* (ref. 7). On the maternal chromosome, the *H19* gene uses the enhancers; on the paternal chromosome, the enhancers are used by *Igf2* (Fig. 1).

But why are the maternal *Igf2* and the paternal *H19* genes silenced? This is where methylation comes in. Methylation can occur at the promoter regions of genes, necessary for the initiation of transcription. The promoter of the paternal *H19* copy is methylated, and this presumably leads to its silencing. But the promoters of the maternal *Igf2* gene are not methylated⁸, so how is this copy silenced? A key insight came after experimental deletion in mice of a paternally methylated region called the imprinting-control region (ICR), located a few kilobases upstream of *H19* (ref. 9; Fig. 1). Deletion of this stretch of DNA from the maternal chromosome (on which this region is unmethylated) results in maternal *H19* and *Igf2* genes both being active. These results raised the possibility that, under normal

circumstances, this region might act as an insulator or 'chromatin boundary element', shielding the maternal *Igf2* promoters from the *H19* enhancers.

Three of the new papers^{1–3} describe the use of different DNA constructs containing the *H19* ICR to test the effect of its position on gene expression. All three groups show that the ICR does indeed have boundary or insulator activity, but apparently no silencing activity. Comparisons of the sequence of these ICRs from mouse, rat and human identified four sites that bind a protein called CTCF^{1,2,4} (Box 1), which is required for boundary function in the chicken β -globin genetic locus¹⁰.

In vitro, CTCF protein bound to all four

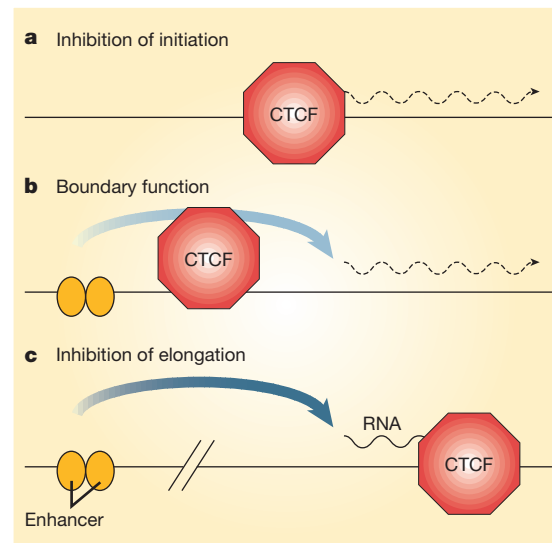
sites on the maternal *H19* ICR^{1,2}. Most important, binding was abolished when the ICR was methylated before the assay was carried out. (Methylation-sensitive binding of protein factors to the ICR has been found previously¹¹, but these proteins may be distinct from CTCF.) Binding of CTCF was also eliminated by mutation of the binding sites, and such mutated fragments no longer showed enhancer-blocking activity. These results establish a strong correlation between methylation and an open (unblocked) boundary, but functional evidence for this link has yet to be obtained. This is not easy — preimposed methylation patterns may not be maintained in these types of assay.

Inheritable mechanisms that control gene expression without affecting gene sequence are said to be epigenetic, and the control of *H19* and *Igf2* expression by the ICR boundary region falls into this category. But how do boundaries work, and what does

Box 1 Activator, repressor and boundary controller

The CTCF protein was originally identified as a ubiquitous repressor of the transcription of chicken genes encoding the oncoprotein c-Myc and the enzyme lysozyme^{14,15}. Most of the functions of CTCF lie in gene repression, but it also has roles in gene activation. It binds to a long stretch of DNA (up to 50 base pairs) by means of its 11-zinc-finger domain. It can use different zinc-fingers to bind to DNA, so the CTCF-binding sites can show considerable sequence divergence. Some of these sites have many CpG (cytosine–guanine) dinucleotides, whereas others have none. Methylation occurs only at CpG sites, so methylation can interfere with the binding of CTCF to only a subset of sites — including the *H19* ICR region^{1,2,4}.

Gene repression can be mediated by CTCF in several ways. These include boundary functions (b in figure), such as blocking the access of enhancer regions to promoters, as shown for *H19* and *Igf2* (refs 1–3). Another function is inhibition of the effects of chromatin-mediated silencing of translocated genes. CTCF can also prevent genes from



being transcribed by removing acetyl groups from histone proteins in chromatin or by inhibiting the transcription-initiation complex (a). Finally, CTCF is found at sites at which the elongation stage of transcription is blocked (c). Different mechanisms of repression may be related; they may complement each other; or they may synergize to result in full repression.

Another protein with many zinc-fingers, *Drosophila* Su(Hw), contains gene-repression domains and is involved in creating a boundary. The clustered binding sites for

Su(Hw) on the 'gypsy' DNA element mediate enhancer blocking in one direction. This is also seen for the *H19* ICR in an assay based on autonomous, self-replicating DNA in vertebrate cells³. The ICR region also confers gene silencing in *Drosophila*, but bidirectionally¹⁶. The *Drosophila* protein that binds to this region has yet to be identified.

Rainer Renkawitz is at the Institute of Genetics, Justus Liebig University, 35392 Giessen, Germany. e-mail: rainer.renkawitz@gen.bio.uni-giessen.de

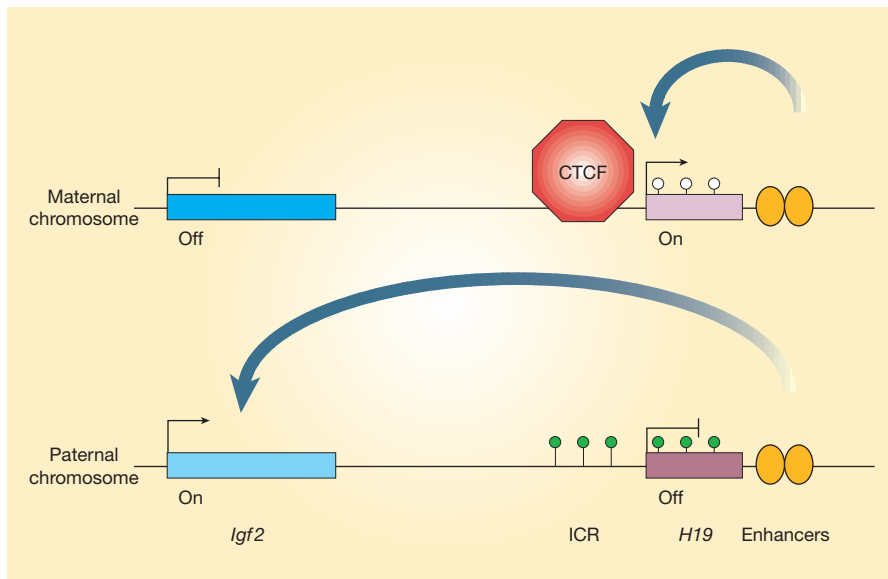


Figure 1 Selective gene silencing by boundaries. *H19* and *Igf2* are genes that are expressed from only the maternal or the paternal chromosome. *H19* is expressed from the maternal chromosome only; *Igf2* is expressed only from the paternal chromosome. The two genes share an enhancer region, located downstream of *H19*. The new papers^{1–4} show that the ICR (imprinting-control region) of *H19* is a boundary element, controlled by DNA methylation. The CTCF protein binds to the unmethylated maternal ICR. This prevents the promoters located in the *Igf2* gene from interacting with the enhancers downstream of the *H19* gene, resulting in transcriptional silencing of *Igf2*. The paternal ICR is methylated (filled circles), preventing binding of CTCF. This allows the enhancers to contact the promoters of the paternal *Igf2*, allowing the gene to be transcribed. The paternal *H19* gene is thought to be silenced by methylation. Differentially methylated regions are also found in *Igf2*, but are not shown here.

CTCF do? First we need to take a step back and look at the way in which enhancers and promoters interact to control gene expression. These two control regions may work together in one of two ways. One way involves the ‘tracking’ of transcription complexes from the enhancers, along the DNA, to the promoter. Alternatively, the DNA may loop round such that the enhancers and promoters interact. The *Igf2* promoters are about 100 kilobases from the enhancers, so the looping model seems more attractive. On the maternal chromosome, either looping or the interaction between the promoters and the enhancers could be disrupted by CTCF (and possibly other factors) binding to the ICR. In addition, binding of CTCF and other factors could have a role in opening the chromatin of the *H19* gene for transcription — CTCF has been implicated in gene activation as well as repression (Box 1).

Indeed, the *H19* ICR seems to have many other functions as well as being a boundary. For example, it is needed for methylation of the paternal *H19* gene⁹, and for silencing of this gene independently of methylation¹². The region seems to contain many functional elements that are involved in regional regulation of methylation, in silencing and in insulating. Other functions may be revealed by further study of these elements and of the protein and chromatin factors that bind to them.

Does the discovery of the epigenetic boundary upstream of *H19* bring to an end

the search for the regulators of *H19* and *Igf2*? Probably not — the fact that deletion of the ICR only partially activates the maternal *Igf2* gene⁹ indicates that other sequences may be involved in keeping the silence. Indeed, deletion of a silencer that is also controlled by methylation and is upstream of *Igf2* also activates the silent *Igf2* allele (M. Constancia *et al.*, unpublished observations). The imprinting arsenal includes promoters, enhancers, antisense RNA transcripts¹³, silencers and chromatin boundaries. What is fascinating is that these sequences have come under epigenetic control. ■

Wolf Reik and Adele Murrell are in the Laboratory of Developmental Genetics and Imprinting, The Babraham Institute, Cambridge CB2 4AT, UK. e-mail: wolf.reik@bbsrc.ac.uk

1. Bell, A. C. & Felsenfeld, G. *Nature* **405**, 482–485 (2000).
2. Hark, A. T. *et al.* *Nature* **405**, 486–489 (2000).
3. Kanduri, C. *et al.* *Curr. Biol.* **10**, 449–457 (2000).
4. Szabo, P. E., Tang, S. H. E., Rentsendorj, A., Pfeifer, G. P. & Mann, J. *Curr. Biol.* **10**, 607–610 (2000).
5. Surani, M. A. *Cell* **93**, 309–312 (1998).
6. Li, E., Beard, C. & Jaenisch, R. *Nature* **366**, 362–365 (1993).
7. Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L. & Tilghman, S. M. *Genes Dev.* **9**, 2079–2089 (1995).
8. Sasaki, H. *et al.* *Genes Dev.* **6**, 1843–1856 (1992).
9. Thorvaldsen, J. L., Duran, K. L. & Bartolomei, M. S. *Genes Dev.* **12**, 3693–3702 (1998).
10. Bell, A. C., West, A. G. & Felsenfeld, G. *Cell* **98**, 387–396 (1999).
11. Ishihara, K. *et al.* *Genome Res.* **10**, 664–671 (2000).
12. Brenton, J. D. *et al.* *Proc. Natl Acad. Sci. USA* **96**, 9242–9247 (1999).
13. Wutz, A. *et al.* *Nature* **389**, 745–749 (1997).
14. Banihmad, A. *et al.* *Cell* **61**, 505–514 (1990).
15. Lobanov, V. V. *et al.* *Oncogene* **5**, 1743–1753 (1990).
16. Lyko, F. *et al.* *Nature Genet.* **16**, 171–173 (1997).



100 YEARS AGO

Possibly some of your photographer readers may be glad to know that microphotography of sorts is within the reach of all who possess a microscope with suitable substage-condenser and a camera... One of my earliest attempts was to photograph fluid inclusions in quartzes with ordinary sunlight, and rock sections polarised. The only difficulty was that the sun would not keep still, and without a heliostat the work was most troublesome, not to say aggravating. In one case, a mere movement of the condenser-diaphragm made the bubble in the inclusion fly backwards and forwards... With a little device in the double lantern the motion of bubbles in inclusions can be shown on a nine-foot screen. These negatives were taken with a 1/16th immersion, the camera being extended with a brown paper tube, and the extra apparatus did not cost one shilling... While observing the transit of Venus, I thought I would try a photograph. I drilled a hole in the telescope cap for diaphragm; took off the eye-piece and stuffed the telescope into a common camera, with a red cloth to make it light-tight; exposed six negatives with hand exposure on instantaneous plates. Result: four passable negatives and one good one. This quite unlooked-for success was due to some back volumes of *Nature* which propped up the camera. From *Nature* 24 May 1900.

50 YEARS AGO

Twilight in India

This book is well named, for the author has a very dim view of India. In spite of a good deal of interesting and, on the whole, well-informed matter on south Indian castes, the whole book is coloured by an obvious determination to view everything Hindu in the darkest shadow, and all the less creditable aspects of Hinduism, particularly in regard to sex, are enlarged on at the expense of its merits.

There is obviously a good deal of exaggeration in many of the statements made, for the meriah sacrifice is written of — and that in 1949 — as if it continued as a routine ceremonial. Several of the illustrations are borrowed without acknowledgement from Thurston; and, although misprints abound, the fact that “tumeric” is repeatedly used for “turmeric” suggests that possibly it is not the printer who is to blame for “foistered” instead of “foisted” — or perhaps “fostered”. From *Nature* 27 May 1950.

Seismology

Inner core takes another turn

Henri-Claude Nataf

Biologists have developed amazing tools to image blood flow in the human body. Is it possible to map motions in the deepest parts of the Earth using similar techniques? This is more or less what Vidale *et al.*¹ (page 445 of this issue) have done. By measuring minute changes in seismic waves that are back-scattered by small heterogeneities in the inner core, the authors find that it rotates some 0.15° per year faster than the mantle. This is the latest, and most refined, estimate of relative core rotation.

The inner core is a ball of solid iron, 1,200 km in radius, at the centre of the Earth. It is surrounded by the outer core, an ocean of liquid iron 2,300 km thick. Convective motions in this electrically conducting ocean produce Earth's magnetic field through the so-called dynamo mechanism. Large-scale eddies at the surface of the outer core have been identified from the analysis of the slow drift of patches of magnetic field, with typical flow velocities of the order of 0.1° per year. These motions could be the surface expression of cylindrical structures extending across the bulk of the outer core because, in a rapidly rotating system, structures tend to be aligned with the axis of rotation.

In 1988, Jault *et al.*² realized that the variation of angular momentum of the outer core, derived from reconstructed core motions between 1970 and 1988, matches that deduced from the variations of the length of the day; these variations are only a matter of milliseconds, but they contained crucial information about the physics of the core. Nevertheless, pinning down motions deep inside the core by measuring the rotation of the solid inner core could provide valuable information about the dynamo's operation.

How can such measurements be made? The idea is to choose a fixed geometry of the emitter and receivers of seismic waves, to record waves that 'see' the inner core, and to look for some variation over a certain period in the characteristics (say, travel times) of those waves.

The first tests along these lines were performed by Souriau in 1989. The emitters were French nuclear explosions in Polynesia and the receiver was the Warramunga seismic array in Australia. The path corresponded to seismic waves that are reflected from the surface of the inner core (Fig. 1a). Hypothetical wide and deep bumps could affect the travel time of these waves, but no definite variations were found.

The second test was the impressive study of Song and Richards³, who analysed the travel time of waves passing through the

inner core (Fig. 1b). The emitters were earthquakes in the South Sandwich islands and the receiver a seismograph in Alaska. Song and Richards found that waves travelling between these two points were 0.3 of a second faster in 1996 than in 1968. The interpretation is subtle and rests on the observation that propagation of seismic waves is anisotropic in the inner core; that is, it varies according to the direction of propagation. Using a model of the distribution of anisotropy, the authors inferred a total inner-core rotation of 30° between 1968 and 1996, giving a rotation rate of about 1° per year. Earlier this year, however, much of the recorded delay was shown to be due to misestimates of the positions of the emitter earthquakes, putting an upper bound on inner-core rotation of 0.2° per year for that period⁴.

Clearly, a higher-resolution technique was needed. This is precisely what Vidale *et al.*¹ provide. The emitters were two nuclear tests in northern Siberia, the devices being exploded in 1971 and 1974. The receiver was LASA, the Large Aperture Seismic Array, a constellation of hundreds of seismic stations in Montana in the United States (the array was dismantled more than 20 years ago but the data it collected remain available). The key ingredient of this new method, however, is the use of waves that are scattered in the inner core rather than transmitted through it (Fig. 1c).

In a previous study⁵, Vidale and Earle discovered that the inner core is not transparent to seismic waves but is somewhat 'fuzzy', a far-reaching result in its own right. Thanks to the high angular resolving power of the LASA array, they obtained a sort of 'angular speckle' of waves scattered by the inner core. They next wondered whether there was any change of this speckle with time. If the inner core rotates with respect to the mantle, waves scattered by heterogeneities located in the side that is getting closer will arrive earlier, whereas waves coming from the retreating side will arrive later.

This is precisely what Vidale *et al.*¹ found when comparing seismic waves from the Siberian nuclear explosions of 1971 and 1974. The observed ± 0.1 -second 'butterfly' pattern is explained by a rotation of only 0.45° , yielding an average eastward rotation rate of 0.15° per year. By looking at scattered rather than transmitted seismic waves, Vidale *et al.* have improved the resolution of these kinds of investigations by almost two orders of magnitude.

Nonetheless, several issues have to be addressed before the measurements of Vidale *et al.* can be seen as definitive. First, theoretical developments are needed to go beyond one element of the analysis — the Born approximation — used in this study. Second, the effect of possibly inexact estimates of emitter location and of heterogeneities near the source requires further work. Finally, the pattern obtained by Vidale *et al.* seems to imply that the axis of rotation of the inner core is different from the Earth's rotation axis. This result is difficult to reconcile with our understanding of core dynamics and will require careful appraisal. Nonetheless, it looks as if here we have a

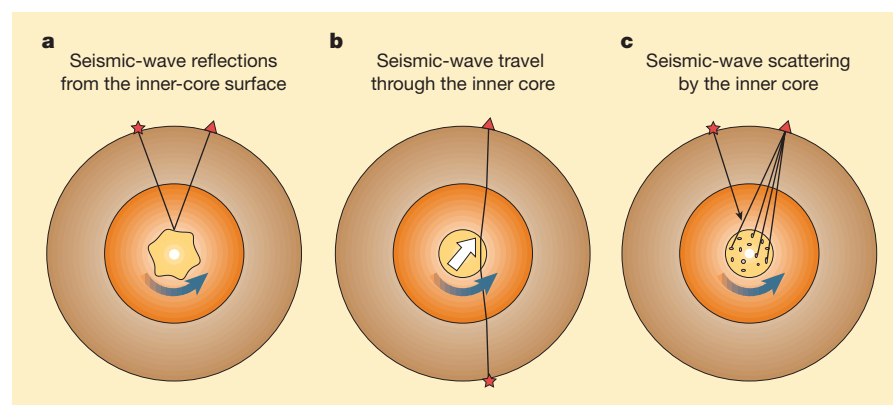


Figure 1 Three ways of detecting differential rotation of the inner core (yellow) with respect to the mantle (brown). Triangles denote a station or array for detecting seismic waves (the receiver); stars indicate an earthquake or a nuclear explosion (the emitter). When an observation is repeated after some time interval, keeping the same geometry, the travel time of the waves may change by a few tenths of a second because the inner core has rotated during that time. a, Souriau analysed waves that bounce off the surface of the inner core. b, Song and Richards³ investigated waves that propagate through the inner core, using its anisotropic axis (white arrow) as a marker. c, The method proposed by Vidale *et al.*¹ relies on waves that are scattered in the inner core. This approach increases the resolution over previous studies by almost two orders of magnitude, and gives an estimate of differential rotation of 0.15° per year.

considerable advance — in principle, the way is open to continuous monitoring of inner-core rotation on periods as short as a year. ■

Henri-Claude Nataf is at the Observatoire des Sciences de l'Univers de Grenoble, CNRS, Université Joseph Fourier, BP 53, 38041 Grenoble, France.

Ocean science

Carbon fixation

Jim Gillon

The total amount of data provided by ocean research is equivalent to that collected daily by meteorologists."

That paraphrased comment, made some years ago, is fast getting out of date. As was evident at a conference* held last month, the upshot of merged research agendas from biology, chemistry and physical oceanography over the past decade is that the balance is being redressed. This is especially so where oceanic carbon is concerned — the central issue for research is of course how the oceans will respond to, and maybe accommodate, the human passion for CO₂ emission and the climate change it is likely to cause.

The greatest uncertainties seem to lie more with the biology than with the physics and chemistry of ocean processes. Maybe this is not surprising. Global-scale oceanography has its roots largely in the physical sciences, so we have a better picture of how physical and chemical parameters — pH, temperature, circulation and so on — can regulate carbon exchange between the ocean and the atmosphere. But it is marine organisms that can really tap into physicochemical carbon cycling and lock up carbon on geological timescales.

Equally, it is the biological response to climate change that seems to be the most difficult to predict. For example, the formation of calcium carbonate shells by certain algae — a process that counterintuitively releases CO₂ — appears to be reduced under conditions of higher levels of atmospheric CO₂ (I. Zonderman, Alfred Wegener Inst., Bremerhaven). So if the balance between carbonate chemistry and photosynthesis shifts in the future, these organisms might switch from being a carbon source to being a carbon sink.

As to the requirements for phytoplankton growth, limiting nutrients such as iron and silicon are known to govern biological productivity in certain oceanic regions. Nutrient supply, whether from the atmosphere or the deep oceans, is bound to change with climate perturbation. But contrary to the more rigid biological assumptions made so far, it seems that algae can modify their nutrient take-up capacity as nutrient avail-

e-mail: henri-claude.nataf@ujf-grenoble.fr

1. Vidale, J. E., Dodge, D. A. & Earle, P. S. *Nature* **405**, 445–448 (2000).
2. Jault, D., Gire, C. & Le Mouél, J.-L. *Nature* **333**, 353–356 (1988).
3. Song, X. & Richards, P. G. *Nature* **382**, 221–224 (1996).
4. Poupinet, G., Souriau, A. & Coutant, O. *Phys. Earth Planet. Inter.* **118**, 77–88 (2000).
5. Vidale, J. E. & Earle, P. S. *Nature* **404**, 273–275 (2000).

ability changes (B. Quéguiner, CNRS, Marseille), with knock-on effects on primary productivity. Such observations call for a more dynamic biology to be incorporated into multi-element biogeochemical models.

Ecological responses further down the food chain also have to be considered. Sinking carbon escapes from surface waters primarily when blooms of larger algae, such as diatoms, outstrip the rate at which the larger grazers can crop the excess. But at the meeting it was clear that we have no answers as to how climate change will affect species assemblages, and hence operation of food chains and carbon export to the ocean depths. Can anything be learned from work on terrestrial systems? Most importantly, experience shows that early conclusions as to the CO₂ response of vegetation often provide little indication of the long-term response because they underestimate plants' ability to acclimatize. Long time-series of data (and the patience to acquire them) are required.

In the meantime there are plenty of other gaps to be filled in, especially in the 'middle ground'. Such areas include the mid-ocean

depths, the 'twilight zone', where sinking carbon is processed (R. Armstrong, SUNY Stony Brook) and physical features at intermediate spatial scales (S. Doney, NCAR, Boulder). These features are on the 10–100-km scale, and are too large to be investigated with shipboard measurements but too fine to be resolved with global approaches such as satellite observation or modelling.

Ocean biogeochemistry models are likely to be in for a shake-up when all these considerations are taken into account. This is no academic matter — the next generation of models will produce the predictions that shape future policy on carbon usage. There are also data to come from the whole-Earth approach, linking up ocean, land and atmosphere, to help understand and quantify the carbon cycle, and from impending further experiments (with nutrient fertilization, for instance). Finally, the new Earth-observing satellites, such as Terra, can even provide information on the physiology of marine plankton as well as its abundance.

The message from the meeting is that the future course of carbon-fuelled research is set fair. But should governments and grant-giving bodies ask where it is heading? Is there still the hope that scientists will show that the Earth may be able to save itself in some Gaiaesque feat of self-sustainability? As the error bars come down on predictions of how the carbon cycle will react to climate change, perhaps policy-makers and industry can move on and face the reality of a different world. Then again, with the prospect of carbon becoming a tradeable commodity, maybe the governmental push to track its every movement is just a sign of a very thorough market-research campaign. ■

Jim Gillon is an associate editor at Nature.

Applied mathematics

The power of design

Mark Newman

The power law is a distinctive experimental signature seen in a wide variety of complex systems. In economics it goes by the name 'fat tails', in physics it is referred to as 'critical fluctuations', in computer science and biology it is 'the edge of chaos', and in demographics and linguistics it is called 'Zipf's law'.

Writing in *Physical Review Letters*¹ and elsewhere², Jean Carlson and John Doyle propose a theory that could help to explain the appearance of power laws in these many different areas. They suggest that power-law distributions, as well as several other features of many complex systems including robustness to perturbations and sensitivity to structural flaws, may be the result of the

design or evolution of systems for optimal behaviour. They call their theory highly optimized tolerance.

A power law is any function of the form $f(x) \propto x^\alpha$, where x is some quantity you are interested in, and α is a constant, usually negative. Many distributions of observed quantities have this power-law form in their tails — that is, for large values of x . Thus, for example, the standardized price returns on individual stocks or stock indices in a stock market have a distribution that falls off approximately as x^{-3} for large returns³; the distribution of population sizes goes as x^{-2} for large cities⁴; the distribution of the sizes of meteor impacts on the Moon⁵, of the numbers of species per genus of flowering plants⁶,

*Joint Global Ocean Flux Study (JGOFS) Open Science Conference, Bergen, Norway, 13–17 April 2000.

of the patterns of activity in J. H. Conway's famous Game of Life⁷, and even of the frequency of citation of scientific papers⁸, all follow power laws.

Is it possible that the common features of all these disparate phenomena could be explained by a single general theory? Some people, notably the proponents of the theory of self-organized criticality, have claimed that they can, but most scientists agree that power-law distributions are the result of many different processes. The distribution of meteor sizes, for example, is almost certainly the product of a random multiplicative or fragmentation process; the behaviour of the Game of Life is an ordinary critical phenomenon; and the distribution of the number of species per genus can be explained by a simple random-walk model. Other power-law-producing mechanisms include the thermal crossing of random energy barriers, systems driven by coherent noise, and the so-called record dynamics. To this toolkit, Carlson and Doyle have now added another, beautiful, idea, which could explain quite a variety of physical phenomena.

Carlson and Doyle first describe their highly optimized tolerance (HOT) theory in the context of a simple 'forest fire' model. This is an attempt to emphasize the similarities and differences between HOT and self-organized criticality, of which the self-organizing forest fire is one of the best-known examples⁹. Imagine then a forest that is managed by a forester who wants to grow as many trees as possible. The principal bane of this forester's life is fire; fires start in the forest with moderate frequency and can destroy large numbers of trees if left unchecked. So the forester cuts fire-breaks to prevent the spread of fire. What is the best way to place these fire-breaks to minimize the damage? If fires are started by sparks which land uniformly at random everywhere in the forest, then the solution is simple — cut the forest into equally sized chunks. However, if there are more sparks in some areas than others, it turns out that the average damage done by a fire is minimized by cutting the forest into chunks whose sizes vary in inverse proportion to the rate at which sparks land in that area.

Carlson and Doyle show that if you take this result and use it to work out the distribution of the sizes of fires, you get a distribution that follows a power law for a wide variety of choices of the distribution of sparks. Thus a power law is generated by the actions of an external agent (the forester) aiming to optimize the behaviour of a system (the forest).

But the HOT mechanism does more than this. If the forest is placed on a regular grid for simplicity, with trees positioned so as to minimize the average damage done by a fire, then the optimal configuration

is one in which the trees are arranged in blocks with discrete fire-breaks between them. So the model tells you not only what the ideal arrangement of fire-breaks is, but also that the best way to control fires is to build breaks. This 'cellular' structure is one of the characteristic features of HOT systems.

Another feature of HOT systems is their sensitivity to unexpected perturbations and design flaws. For instance, if the distribution of positions at which fires start changes from the one for which the tree configuration is optimized, it can cause catastrophic damage; the average fires can be much larger in this case than if the trees were uniformly clumped. And if one of the fire-breaks has a flaw — a single fallen tree across the break, for example — this can result in much worse damage than such a small perturbation seems to warrant. Carlson and Doyle point out that these phenomena are well known to engineers who design systems for optimal performance. Highly tuned systems are often sensitive to small imperfections, so engineers commonly design them to be slightly suboptimal to avoid such problems.

This last point is crucial to the HOT picture. Carlson and Doyle have pitched HOT not only as a mechanism for generating power-law distributions, but also as a way of quantifying ideas about designed systems which are well known (anecdotally) in engineering, in a way that makes them comprehensible and useful to the scientific

community. In this sense, Carlson and Doyle's paper¹ succeeds very well, constructing a theory of 'robust yet fragile' designs in a language that will be comfortably familiar to many of us. However, their general conceptual approach, and also the presentation of the idea in terms of abstract systems such as forest-fire models, is inevitably going to lead people to ask what real applications HOT has. Is HOT an answer looking for a question?

It seems not, and to prove it Carlson and Doyle have applied their ideas to a variety of real-world systems. In new work, as yet unpublished, they show how HOT can explain data on real forest fires, electrical power failures and Internet traffic. They claim that further examples are easy to find. We won't have to look hard. And if they are right, HOT could be one of the most important additions to the theory of complex systems in recent years. ■

Mark Newman is at the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA.

e-mail: mark@santafe.edu

1. Carlson, J. M. & Doyle, J. *Phys. Rev. Lett.* **84**, 2529–2532 (2000).
2. Carlson, J. M. & Doyle, J. *Phys. Rev. E* **60**, 1412–1427 (1999).
3. Mantegna, R. N. & Stanley, H. E. *Nature* **376**, 46–49 (1995).
4. Zipf, G. K. *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Reading, 1949).
5. Neukum, G. & Ivanov, B. A. in *Hazards Due to Comets and Asteroids* (ed. Gehrels, T.) 359–416 (Univ. Arizona Press, Tucson, 1994).
6. Willis, J. C. *Age and Area* (Cambridge Univ. Press, 1922).
7. Bak, P., Chen, K. & Creutz, M. *Nature* **342**, 780 (1989).
8. Redner, S. *Eur. Phys. J. B* **4**, 131–134 (1998).
9. Drossel, B. & Schwabl, F. *Phys. Rev. Lett.* **69**, 1629–1632 (1992).

Neurobiology

Self-regulating synapses

Christine R. Rose and Arthur Konnerth

Information transfer between neurons is achieved mainly through chemical synapses. The presynaptic neuron releases a neurotransmitter into the synaptic cleft, and the transmitter binds to synaptic receptors on the postsynaptic cell. This generally leads to the opening of ion channels in the postsynaptic membrane and an alteration in the electrical properties of the postsynaptic neuron. The transmission properties of chemical synapses are controlled by neuronal activity. This activity-dependent regulation of synaptic plasticity is thought to represent the cellular basis for the development of neural circuitry and to underlie learning and memory. Hence the importance of the paper on page 454 of this issue¹, in which Liu and Cull-Candy reveal a mechanism for the regulation of synapses in the brain that are responsive to the neurotransmitter glutamate.

Over the past decade, several mecha-

nisms have been described by which neuronal activity alters the transmission properties of synapses that use glutamate (see ref. 2 for a review). Most of these mechanisms are based on an activity-induced alteration in the number and/or functional properties of one type of glutamate receptor — the so-called AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionate) receptor. AMPA receptors are the main contributors to the excitatory postsynaptic current that can be generated at low frequencies of stimulation. Increased neuronal activity can induce the phosphorylation of AMPA receptors, leading to an increase in the flow of ions through AMPA channels³, or it can induce AMPA-receptor-type activity at synapses that were not previously responsive⁴. The latter process is likely to be mediated by activity-regulated insertion of AMPA receptors at the synapse^{5,6}.

The activity-induced alteration of AMPA-

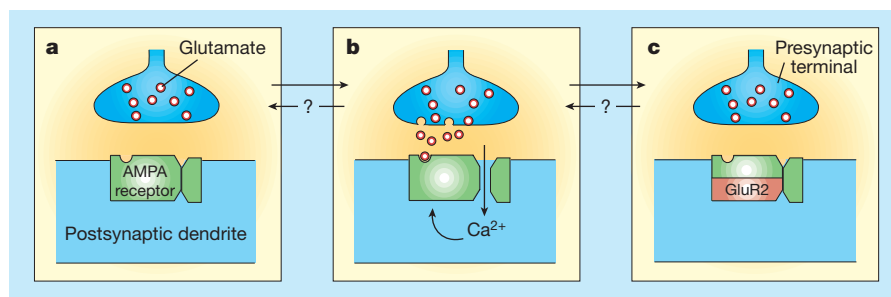


Figure 1 A new way of regulating synaptic activity. Liu and Cull-Candy¹ have found that the composition of AMPA receptors (a subtype of receptors for the neurotransmitter glutamate) in cerebellar stellate interneurons can change in response to neuronal activity. These receptors regulate their own function, perhaps as follows. **a**, Under control conditions, synapses are characterized by a high proportion of calcium-permeable AMPA receptors that lack the subunit GluR2. **b**, Release of glutamate from the presynaptic neuron and its binding to AMPA receptors of the postsynaptic neuron leads to calcium influx through the AMPA receptors. The resulting increase in intracellular calcium concentration induces an alteration in the subunit composition of the AMPA receptors themselves (**c**), resulting in the expression of GluR2-containing receptors with low calcium permeability.

receptor function usually requires an influx of calcium ions into the neurons. How is this increase in intracellular calcium concentration achieved? Many neurons express AMPA receptors that exhibit very low permeability to calcium, because the receptors contain a specific subunit, GluR2 (ref. 7). Calcium influx into these neurons can be achieved by activation of another type of glutamate receptor, the NMDA (N-methyl-D-aspartate) receptor, or through voltage-gated calcium channels in the plasma membrane. In contrast, AMPA receptors that lack GluR2 exhibit relatively high calcium permeability. Cells expressing a high proportion of these receptors can use this route for calcium entry.

Liu and Cull-Candy¹ now describe a mechanism of synaptic plasticity based, unusually, on changes in the composition and function of AMPA receptors. They provide evidence that activity regulates the molecular composition of AMPA receptors in postsynaptic stellate interneurons of the cerebellum. The authors use a variety of pharmacological tools to show that these interneurons do not use NMDA receptors. Instead, the synapses are characterized by a high proportion of calcium-permeable, GluR2-lacking AMPA receptors. By contrast, at the cell bodies of the interneurons (that is, at extrasynaptic regions), Liu and Cull-Candy find mainly GluR2-containing AMPA receptors, with low calcium permeability.

What are the mechanisms responsible for the targeting of the different AMPA-receptor subunits to different subcompartments of these neurons? Liu and Cull-Candy tackled this problem by applying high-frequency stimulation to the presynaptic axons. Surprisingly, within just 15 to 30 minutes, the properties of the AMPA-receptor-mediated currents in the postsynaptic interneurons changed strikingly, indicating

the inclusion of GluR2-containing receptors. The results convincingly show the rapid, activity-dependent change in function at these synapses.

Liu and Cull-Candy next went on to show that an increase in the level of intracellular calcium was necessary for the alteration in receptor composition. They then analysed the different routes of calcium entry into the cell. They discovered that calcium influx through the synaptic AMPA receptors themselves was sufficient to produce a rise in intracellular calcium that resulted in alteration of the function of these receptors. By contrast, the function of extrasynaptic receptors on the cell bodies was determined mostly by calcium influx through N-type calcium channels during action potentials. So, at some synapses in the brain, receptor-mediated calcium influx can regulate the subunit composition — and thereby the calcium permeability and the electrical properties — of the very same receptors. This provides an activity-dependent feedback mechanism controlling the properties of synaptic transmission.

What might be the physiological function of this self-regulating mechanism? At present, this question is difficult to answer. Perhaps these synapses function in two ways

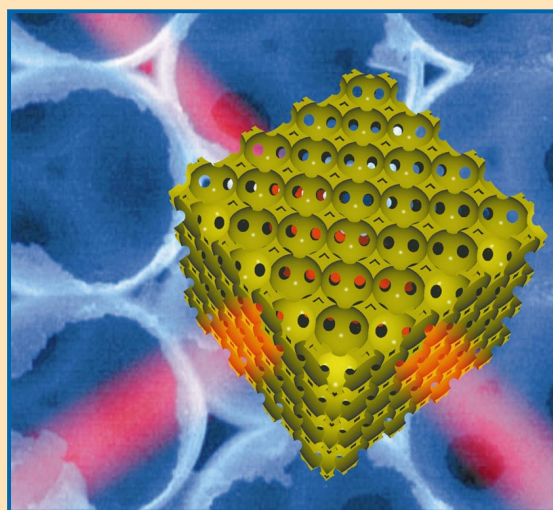
Photonics

Opal appeal

Photons are the best messengers for really quick information transmission. But creating circuitry that can process light with the versatility of electronic silicon chips is a big technological challenge. The most promising approach relies on fabricated structures known as 'photonic crystals'. These materials have carefully tailored properties that make them opaque to selected wavelengths — providing a means to control the routing and transmission of optical signals.

Elsewhere in this issue, Alvaro Blanco and colleagues (*Nature* **405**, 437–440; 2000) describe a simple and cheap method for the large-scale manufacture of three-dimensional silicon photonic crystals. The picture here shows a computer representation of the structures, superimposed on a real image from a scanning electron microscope. Because the crystals are made of silicon, they should be easy to integrate with conventional electronic circuitry.

How is this complex



structure produced? The first step is to construct a three-dimensional template that will ultimately be discarded. It consists of a form of artificial opal: uniformly sized silica spheres, up to a micrometre in diameter, grown from seed particles. Application of heat causes small bridges to develop between the spheres, resulting in a connected structure. Next, gases containing silicon are allowed to infiltrate the template. The voids in between the spheres are gradually filled up as the silicon crystallizes. The final step is to dissolve

away the silica sphere-and-bridge template using an acid etching process.

Theoretical calculations predict that the resulting structure should block wavelengths around 1.5 μm — perfect for manipulating the infrared signals used in fibre-optic communications. The measured reflectance and transmission spectra of the photonic crystals confirmed this behaviour. Equally importantly, the crystals are unlikely to absorb valuable signal power, as bulk silicon only soaks up higher energies. **Karen Southwell**

(Fig. 1). In control conditions, the complex synaptic response would consist of both the electrical response and a transient calcium signal. Activity-induced insertion of GluR2-containing AMPA receptors into the postsynaptic membrane would switch the synapse to a second mode, in which the calcium signal is suppressed because of the reduced calcium entry. Another plausible, but equally speculative, role for the switch in subunit composition is that it serves as a mechanism to scale transient changes in postsynaptic calcium levels⁸. Or it could have a purely neuroprotective function; calcium entry through AMPA receptors has been implicated in the neurodegeneration associated with ischaemia (reduced blood flow) in the brain and epilepsy.

Before we can completely understand this process and its implications, it will be important to determine the calcium-dependent events that lead to the insertion of a GluR2-containing receptor complex into the postsynaptic membrane. Also interesting is how GluR2 might be removed to reset the system and to re-establish the functional properties found in control conditions. The

answers to these questions may come with analysis of the amplitude and kinetics of the transient changes in intracellular calcium that induce these effects. In fact, different patterns of such calcium transients may either upregulate or downregulate synaptic function through activation of calcium-permeable AMPA receptors in other neurons^{9,10}. The self-regulation of AMPA receptors revealed by Liu and Cull-Candy¹ has added yet another piece to the puzzle of synaptic plasticity. ■

Christine R. Rose and Arthur Konnerth are at the Institute of Physiology, Technical University of Munich, Biedersteiner Strasse 29, 80802 Munich, Germany.

e-mail: konnerth@physiol.med.tu-muenchen.de

1. Liu, S.-Q. & Cull-Candy, S. G. *Nature* **405**, 454–458 (2000).
2. Malenka, R. C. & Nicol, R. A. *Science* **285**, 1870–1874 (1999).
3. Benke, T. A. *Nature* **393**, 793–797 (1998).
4. Durand, G. M., Kovalchuk, Y. & Konnerth, A. *Nature* **381**, 71–75 (1996).
5. Hayashi, Y. *et al.* *Science* **287**, 2262–2267 (2000).
6. Song, I. *et al.* *Neuron* **21**, 393–400 (1998).
7. Burnashev, N. *et al.* *Neuron* **8**, 189–198 (1992).
8. Turrigiano, G. G. *et al.* *Nature* **391**, 892–896 (1998).
9. Gu, J. G. *et al.* *Nature* **381**, 793–796 (1996).
10. Laezza, F., Doherty, J. J. & Dingledine, R. *Science* **285**, 1411–1414 (1999).

DNA repair

The bases for Cockayne syndrome

Philip C. Hanawalt

Cockayne syndrome is a rare human hereditary disease, characterized by growth failure, deficient neurological development and severe sensitivity to sunlight. It can arise from mutations in any one of five genes. The protein products of these genes are involved in different aspects of the repair of damaged DNA, and it has been far from clear how all these different mutations result in the same syndrome. Le Page and colleagues¹, writing in *Cell*, now provide important clues to the answer. It seems that the common problem in cells from patients with Cockayne syndrome is a failure to repair oxidation-induced damage to DNA bases, specifically in the strands of DNA that are being transcribed into RNA.

Different DNA-repair pathways operate on different types of DNA lesions. Nucleotide-excision repair (NER), for example, is a ubiquitous cellular process by which short, single-stranded DNA segments, containing damaged nucleotides, are removed from duplex DNA. The gaps are then filled in by repair DNA synthesis, using the intact strand as a template (see ref. 2 for a review). Defects in NER underlie the hereditary disease xeroderma pigmentosum. This disease — like Cockayne syndrome — is characterized by severe sensitivity to sunlight. However, patients with xeroderma pigmentosum

are several thousand times more likely than Cockayne syndrome patients to develop cancer in exposed areas of skin. Otherwise, mutations in most of the seven XP (xeroderma pigmentosum) genes needed for NER of photoproducts in DNA do not usually pose serious health problems. Another pathway, termed base-excision repair (BER), operates on the damage to bases produced by reactive oxygen species, ionizing radiation and some alkylating agents, as well as certain inappropriate bases (such as uracil) in DNA.

Yet another process, transcription-coupled repair (TCR), deals with a variety of DNA lesions that are thought to arrest the transcription of genes³. This process has been considered for some time to be a type of NER. If the ultraviolet wavelengths of sunlight cause damage to the strand of a DNA duplex that is being transcribed into RNA, TCR solves the problem. By contrast, lesions throughout the genome — including ultraviolet-light-induced damage to the non-transcribed strands of expressed genes — are repaired by global genomic NER.

Mutations in either of two non-essential genes — CSA or CSB — result in defective TCR, and are the genetic defect in over 90% of Cockayne syndrome patients. Certain mutations in XP genes also underlie a small number of Cockayne syndrome cases. Two

of these genes, *XPB* and *XPD*, encode components of the general transcription factor TFIIH. This complex is needed to open up the DNA strands in preparation for the enzyme RNA polymerase II to begin transcription. It also opens up regions that include a DNA lesion, allowing NER to take place. The third XP gene so involved is *XPG*, which encodes a protein required to make the first of the two incisions in the DNA strand needed for NER.

Thus all the mutations that cause Cockayne syndrome have in common the property that they eliminate TCR of ultraviolet-damaged DNA. This explains the sensitivity to sunlight, but what about the developmental defects, which are unlikely to result from ultraviolet damage to DNA? Might the basis for these defects lie in the defective NER of similar damage caused by other agents? This seems unlikely, as mutations in the *XPA* gene (which is involved in lesion recognition) that totally eliminate both global genomic NER and the TCR of such damage do not result in Cockayne syndrome. A second hypothesis is that the disease is a 'transcription syndrome', in which certain groups of genes are deficiently expressed⁴. In this model, the mutations in CSA and CSB, like those in *XPB* and *XPD*, are envisaged to have direct effects on transcription itself. But it is hard to explain how the mutations in *XPG* affect transcription.

So what is the basis for Cockayne syndrome? Getting to the answer requires rethinking the relationship between TCR and NER. It seems that, far from being a subpathway of NER, TCR may in fact act upstream of both nucleotide- and base-excision repair.

A first step along the way to this answer was provided by the report⁵ that DNA damage produced by ionizing radiation (not ultraviolet radiation) is subject to TCR in normal human cells and in *XPA* mutant cells, but not in *CSB* mutant cells. DNA damage produced by ionizing radiation is generally thought to be remedied by BER, not NER. In addition, TCR of an oxidized base, thymine glycol, has been shown to be defective as a result of *XPG* mutations that result in Cockayne syndrome⁶, but not of other *XPG* mutations that result only in the symptoms of xeroderma pigmentosum⁷. Repair of thymine glycol is also achieved by BER. So, TCR can be linked to BER as well as to NER. These results are further evidence that Cockayne syndrome might result from defective TCR of oxidative lesions.

Le Page *et al.*¹ have now finished testing the hypothesis that all patients with Cockayne syndrome — no matter which gene is mutated — should be deficient in the TCR of oxidative lesions, whatever their nature. Their results firmly establish that both the *XPG* protein and TFIIH have

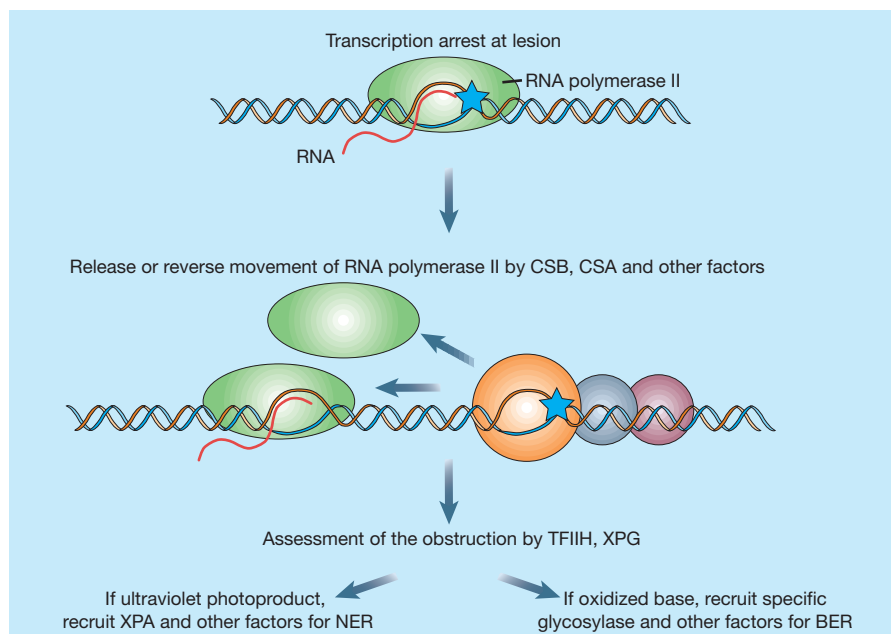


Figure 1 The DNA-repair pathways meet. The results of Le Page *et al.*¹ lead to the following model for the repair of lesions in the transcribed strands of expressed genes. Top, RNA polymerase II transcribes DNA into RNA. If RNA polymerase II meets a lesion (asterisk) in the DNA strand, it will stop. Centre, repair of the lesion and further transcription is blocked until the polymerase is either released from the DNA or backed away from the lesion. Both TFIIH and XPG are then required before the lesions are repaired, perhaps to determine the nature of the obstruction. Bottom, enzymes of either the nucleotide-excision-repair (NER) or the base-excision-repair (BER) pathways are recruited, to remove the offending lesion and allow transcription to resume. The genes that can be mutated in Cockayne syndrome are required to remove the stalled polymerase, and to assess the nature of the damage.

different roles in the TCR of oxidative damage and in NER.

Le Page *et al.* also turned their attention to another prominent oxidative lesion, 8-oxo-guanine. Using a shuttle vector, they introduced DNA containing 8-oxo-guanine residues into several human cell strains. They found that 8-oxo-guanine, like thymine glycol, is subject to TCR, but not significantly to NER — so it must be subject to BER instead. They also showed that 8-oxo-guanine (or 8-oxo-guanine plus a ligand) is an obstacle to the movement of RNA polymerase II, and that the arrested polymerase must be released to give the repair enzymes access to the site. So, not only is TCR missing in Cockayne syndrome, but the blocked RNA polymerase II prevents lesion recognition and repair by the global BER pathway as well.

Instead of being considered a subpathway of NER, it now seems that TCR is a process needed before either the BER or the NER pathway can repair oxidative damage to genes that are being transcribed, in which RNA polymerase is stalled at a lesion. TCR might be considered an 'obstacle-recognition' step, in which TFIIH and XPG may work together to assess the nature of the obstruction (Fig. 1). This obstruction might be removed only when the arrested RNA polymerase II has been moved away, which may be achieved by CSA and CSB, amongst other factors. TFIIH and XPG might then

recruit the enzymes needed for BER or NER, as appropriate.

It is clear that an impasse provided by an arrested RNA polymerase II would severely affect transcription. In that sense, Cockayne syndrome could, indeed, be considered a 'transcription syndrome'. But the arrest of transcription also provides a strong signal for a cell-death (apoptosis) pathway⁸. So, Cockayne syndrome could also be characterized as a disease of excessive cell death by apoptosis — a disease that would affect rapidly metabolizing cells, such as neurons, that generate high levels of reactive oxygen species. The apoptosis model could also, in principle, explain the problems of stunted growth and neurological deterioration. It might also explain why Cockayne syndrome patients are not prone to skin cancer, even in the face of severe sunlight sensitivity — after all, dead cells do not form tumours. ■

Philip C. Hanawalt is in the Department of Biological Sciences, Stanford University, Stanford, California 94305-5020, USA.

e-mail: hanawalt@stanford.edu

1. Le Page, F. *et al.* *Cell* **101**, 159–171 (2000).
2. Lindahl, T. & Wood, R. D. *Science* **285**, 1897–1905 (1999).
3. Tornaletti, S. & Hanawalt, P. C. *Biochimie* **81**, 139–148 (1999).
4. Bootsma, D. & Hoeijmakers, J. H. J. *Nature* **363**, 114–115 (1993).
5. Leadon, S. A. & Cooper, P. K. *Proc. Natl Acad. Sci. USA* **90**, 10499–10503 (1993).
6. Cooper, P. K. *et al.* *Science* **275**, 990–993 (1999).
7. Nussipikel, T. *et al.* *Proc. Natl Acad. Sci. USA* **94**, 3116–3121 (1997).
8. Ljungman, M. & Zhang, F. *Oncogene* **13**, 823–831 (1996).

Daedalus

The depths of madness

Protein molecules, folded so subtly, work all the enzymatic wonders of biochemistry. Heat can unfold and denature them — as all cooks know. But so can pressure. Raw egg white is opacified and hardened by a few thousand atmospheres (though it still tastes raw). Daedalus is now pondering the implications for deep-sea creatures.

A crucial problem for such creatures is to know how deep they are. Some rely on their eyes — daylight from the surface fades out with depth. Others have a swim-bladder, whose gas is compressed by the rising hydrostatic pressure as its owner sinks. But below 1 km or more, neither scheme works well. Daedalus reckons that whales, squid and other creatures who repeatedly patrol a wide vertical range exploit pressure-dependent protein folding. Some 'barometric enzyme' in their make-up changes its reactivity reversibly with pressure to register their depth.

So DREADCO trawlers are now lowering pressurized traps into the oceans, to snare abyssal creatures and haul them to the surface still under pressure. They will then be transferred to a pressurized aquarium for study. They, or perhaps their symbiotic and gut organisms, should harbour novel barometric enzymes which could throw new light on the protein-folding problem.

But this costly research programme has a greater goal. Daedalus argues that all the proteins of an abyssal creature must unfold somewhat in the deeps. And even if they fold again when their owner rises, very likely at least one molecule will do so wrongly. It will form a prion — and once one has formed, it will catalyse production of more prions the next time the creature dives. In the brain, proliferating prions pose a lethal threat. What protects a squid from 'mad squid disease'?

Whatever it is, Daedalus wants to know. For we also are vulnerable to misfolded proteins, not only as the brain plaques of Alzheimer's disease and CJD, but in the slow advance of wrinkled skin and hardened eye lenses and arteries. Somewhere in abyssal biochemistry, he hopes, is the crucial refolding reagent that prevents a protein from prionizing or forming a plaque, or even hauls it back from that doom. It could be an elixir of health and youth for all mankind. Daedalus cannot guess how it may work. But his team is on the look-out for any separated fraction that can unboil an egg. **David Jones**

The Further Inventions of Daedalus is published by Oxford University Press.

Wireless capsule endoscopy

The discomfort of internal gastrointestinal examination may soon be a thing of the past.

We have developed a new type of endoscopy, which for the first time allows painless endoscopic imaging of the whole of the small bowel. This procedure involves a wireless capsule endoscope and we describe here its successful testing in humans.

The invention of fibre-optic endoscopy¹ made visualization of the whole stomach, upper small bowel and colon possible. The procedures used to examine these (gastroscopy, small-bowel endoscopy and colonoscopy, respectively) cause discomfort because they require flexible, relatively wide cables to be pushed into the bowel — these cables carry light by fibre-optic bundles, power and video signals. Small-bowel endoscopy in particular is constrained by problems of discomfort and limitations of how far enteroscopes can be advanced into the small bowel. There is a clinical need for improved methods of examining the small bowel and colon, especially in patients with recurrent gastrointestinal bleeding.

The invention of the transistor made it possible to design swallowable electronic radio-telemetry capsules for the study of gastrointestinal physiological parameters. These capsules were first reported in the 1950s and were used to measure temperature², pressure^{2,3} and pH^{3,4}. We have developed and tested a new type of video-telemetry capsule endoscope that is small enough to be swallowed (11 × 30 mm) and has no external wires, fibre-optic bundles or cables. By using a lens of short focal length, images are obtained as the optical window of the capsule sweeps past the gut wall, without requiring air inflation of the gut lumen. The capsule endoscope is propelled by peristalsis through the gastrointestinal tract and does not require a pushing force to propel it through the bowel.

The video images are transmitted using UHF-band radio-telemetry to aerials taped to the body which allow image capture, and the signal strength is used to calculate the position of the capsule in the body (see Supplementary Information); the images are stored on a portable recorder. This system allows more than 5 hours of continuous recording. The patient need not be confined to a hospital environment during the examination and is free to continue his or her daily routine.

The design of the video capsule was made possible by progress in the performance of three technologies: complementary metal oxide silicon (CMOS) image sensors, application-specific integrated circuit (ASIC) devices, and white-light-

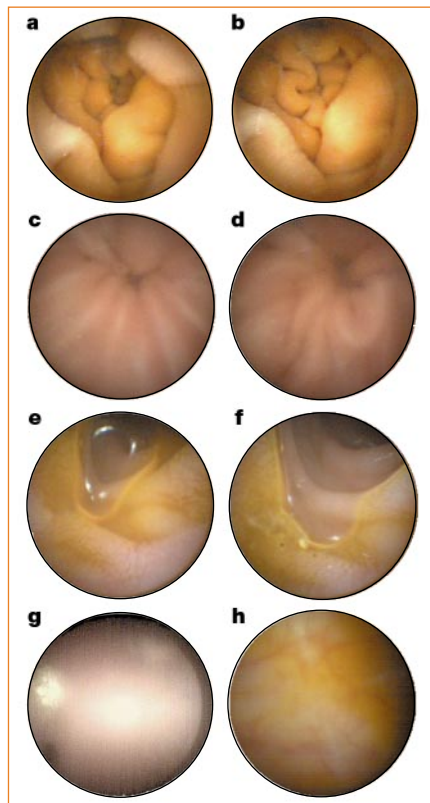


Figure 1 Samples of images of the small bowel acquired by the capsule endoscope during human *in vivo* studies. **a,b**, Gastric folds in the body of the stomach; **c,d**, villous pattern of the small bowel enhanced by the presence of a little water and an air bubble in the lumen; **e,f**, airless images of normal jejunum, viewed with the lumen closed in front of the optical dome of the capsule; **g,h**, views of the terminal ileum.

emitting diode (LED) illumination. Novel optical design, better energy management and overall system design were also important in creating the capsule.

The addition of a buffer amplifier on each pixel reduced the output noise that

was initially associated with CMOS image sensors and has allowed CMOS chips to achieve an image quality comparable to those of charge-coupled device image sensors⁵, but using much less power.

Advances in ASIC design allowed the integration of a very small video transmitter of sufficient power output, efficiency and bandwidth into the capsule. Synchronous switching of the LEDs, the CMOS sensor and the ASIC transmitter minimize power consumption. By careful design of the optics, we were able to eliminate internal reflections which are a common problem when the illumination and imager are incorporated under the same dome.

With ethical committee approval, the first studies were performed on ten normal human volunteers. The capsule was easily swallowed and caused no discomfort. Propelled by peristalsis (see Supplementary Information), it successfully transmitted video images (Fig. 1) from the stomach, small bowel and caecum (mean gastric transit time was 80 min, range 17–280 min; mean small-bowel transit time was 90 min, range 45–140 min; mouth-to-evacuation time was 24 h, range 10–48 h). High-quality images were received throughout the video transmissions, lasting up to 6 hours.

Gavriel Iddan*, Gavriel Meron*, Arkady Glukhovskiy*, Paul Swain†

*Given Imaging Ltd, Building 7, New Industrial Park, Yoqneam 20692, Israel

†Royal London Hospital, Whitechapel, London E1 1BB, UK

1. Hopkins, H. H. & Kapany, N. S. *Nature* **173**, 39–41 (1954).
2. Zworkin, V. K. *Nature* **179**, 898 (1957).
3. Mackay, R. S. & Jacobson, B. *Nature* **179**, 1239–1240 (1957).
4. Noller, H. G. *Deutsche Med. Wsch.* **85**, 1707 (1960).
5. Fossum, E. R. *Proc. SPIE* **1900**, 2–14 (1993).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of *Nature*.

Cell biology

Non-thermal heat-shock response to microwaves

Exposure limits set for microwave radiation assume that any biological effects result from tissue heating¹: non-thermal effects have been reported but remain controversial. We show here that prolonged exposure to low-intensity microwave fields can induce heat-shock responses in the soil nematode *Caenorhabditis elegans*. This effect appears to be non-thermal, suggesting that current exposure limits set

for microwave equipment may need to be reconsidered.

Heat-shock proteins (HSPs) are induced in most organisms by adverse conditions (such as heat or toxicants) that cause damage to cellular proteins, acting as molecular chaperones to rescue damaged proteins². To detect HSP responses, we have pioneered the use of transgenic *C. elegans* strains carrying reporter-gene constructs (encoding β -galactosidase in strain PC72 or green fluorescent protein (GFP) in strain PC161) regulated by homologous *hsp16* heat-shock promoters³. When exposed to diverse stressors at 20–25 °C,

these worms express readily detectable reporter products, whereas controls show minimal expression⁴.

Worms were exposed overnight to continuous-wave microwave radiation at 750 MHz and 0.5 W in the transverse electromagnetic (TEM) cell described previously⁵. Figure 1 shows temperature profiles for reporter expression in both irradiated and control (foil-shielded) worm cultures. In microwave-exposed cultures, expression is comparable to that of controls at 24.0 °C ($P>0.05$), but then rises steeply through 24.5 and 25.0 to 25.5 °C ($P<0.001$). In non-exposed controls, heat-induced reporter expression follows the pattern for HSP16 (ref. 6), increasing sharply only above 27 °C (to a maximum at 30 °C). There is thus a disparity of 3 °C between exposed and control induction profiles.

A thermal explanation for this disparity would require that the exposed worms become 3 °C warmer than controls — or more if only a minority of worms/tissues is affected. We reject this thermal explanation on several grounds, not least the diffusion of heat over 18 hours.

First, no temperature difference is detectable between control and exposed cultures after irradiation⁵. This is also true for concentrated (50% w/v) worm suspensions incubated for 18 h at 25 °C alongside a saline solution alone, under exposed versus control conditions (24.68 ± 0.116 °C s.d., $P=0.28$, for all 16 measurements under four conditions using a sensitive copper-constantan microthermocouple). Temperature differences of 0.5 °C (that is, worms

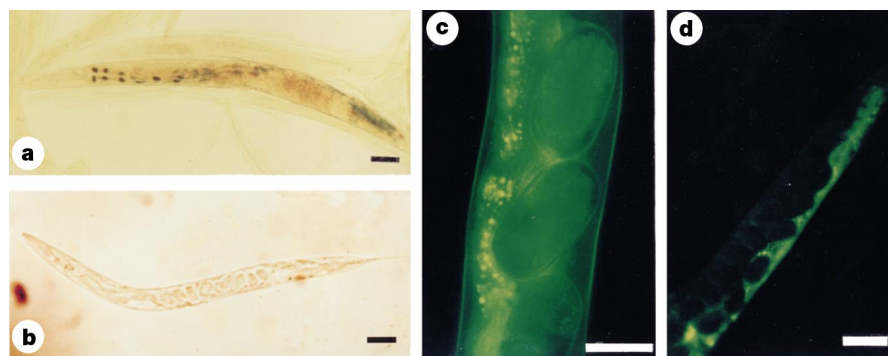


Figure 2 PC72 and PC161 (similar to PC72, but carrying an additional GFP reporter under *hsp16* control) worms were either exposed for 18 h at 25 °C to microwaves (750 MHz, 0.5 W) or kept as 25 °C controls, then reporter expression was localized *in situ* by staining with X-gal (PC72) or viewing under ultraviolet light on a fluorescence microscope (PC161). **a**, Exposed PC72 worm, showing nuclear staining for β-galactosidase throughout the gut; **b**, typical PC72 worm under control conditions: no observable staining; **c**, exposed PC161 worm, showing GFP fluorescence throughout ovoid embryos; **d**, typical control PC161 worm, showing yellowish gut autofluorescence (also in **c**) but no GFP fluorescence in embryos. Note that many worms in **a** and **c** show little reporter expression. Scale bars, 50 μm.

1 °C warmer than the saline) would have been easily detectable in this experiment.

Second, *in situ* detection of reporter products shows that *lacZ* is expressed throughout the gut in PC72 worms (Fig. 2a,b), and also that GFP is expressed in many embryos within adult PC161 worms (Fig. 2c,d). These expression sites together constitute about half of worm tissues.

Third, the field at the centre of our TEM cell is 45 V m⁻¹, and the measured permittivity of concentrated worm suspensions (at 615 MHz) gives a conductivity of about 0.48 Ω⁻¹ m⁻¹. The calculated specific absorption rate (SAR) is only 0.001 W kg⁻¹, which is much less than published values⁷ for mobile phones (0.02–1.0 W kg⁻¹). Mobile-phone manufacturers claim that SARs in this range are insufficient to cause measurable tissue heating within the human head, and we are not disputing this.

We suggest instead that the induction of heat-shock proteins described here could involve non-thermal mechanisms. These could include microwave disruption of the weak bonds that maintain the active folded forms of proteins; enhanced production of reactive oxygen species (known to be inducers of HSPs⁸); or interference with cell-signalling pathways that affect HSP induction (by heat-shock-factor activation). All these mechanisms are testable using the functional genomic tools that are available in *C. elegans*. Because of the universality of the heat-shock response², a similar non-thermal induction might also occur in human tissues exposed to microwaves, a possibility that needs investigation.

David de Pomerai*, **Clare Daniells***, **Helen David***, **Joanna Allan***, **Ian Duce***, **Mohammed Mutwakil***, **David Thomas†**, **Phillip Sewell†**, **John Tattersall‡**, **Don Jones§**, **Peter Candido§**

*Molecular Toxicology Division, School of Biological Sciences and †School of Electrical and Electronic Engineering, University of Nottingham, Nottingham NG7 2RD, UK

‡Medical Countermeasures, CBD Porton Down, Salisbury, Wiltshire SP4 0JQ, UK

§Department of Biochemistry and Molecular Biology, University of British Columbia, 2146 Health Sciences Mall, Vancouver V6T 1Z3, Canada

1. ANSI/IEEE C95.1-1992 American National Standard-Safety Levels with Respect to Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz (IEEE, New York, 1992).
2. Parsell, D. & Lindquist, S. *Annu. Rev. Genet.* **27**, 437–496 (1993).
3. Dennis, J., Mutawakil, M., Lowe, K. & de Pomerai, D. *Aquatic Toxicol.* **40**, 37–50 (1997).
4. Candido, P. & Jones, D. *Trends Biotechnol.* **40**, 125–129 (1996).
5. Daniells, C. *et al. Mutat. Res.* **399**, 55–64 (1998).
6. Snutch, T. & Baillie, D. *Can. J. Biochem. Cell Biol.* **61**, 480–487 (1983).
7. Gandhi, O., Lazzi, G. & Furse, C. *IEEE Trans. Microwave Theor. Tech.* **44**, 1884–1897 (1996).
8. Nishizawa, J. *et al. Circulation* **99**, 934–941 (1999).
9. Williams, P. & Dusenbery, D. *Environ. Toxicol. Chem.* **9**, 1285–1290 (1990).
10. Jewitt, N., Anthony, P., Lowe, K. & de Pomerai, D. *Enzyme Microb. Technol.* **25**, 349–356 (1999).

Structural biology

Proton-powered turbine of a plant motor

ATP synthases are enzymes that can work in two directions to catalyse either the synthesis or breakdown of ATP, and they constitute the smallest rotary motors in biology. The flow of protons propels the rotation¹ of a membrane-spanning complex of identical protein subunits, the number of which determines the efficiency of energy conversion. This proton-powered turbine is predicted to consist of 12 subunits^{2–4}, based on data for *Escherichia coli*⁵. The yeast mitochondrial enzyme, however, has only 10 subunits⁶. We have imaged the ATP synthase from leaf chloroplasts by using atomic force microscopy and, surprisingly, find that its turbine has 14 subunits, arranged in a cylindrical ring.

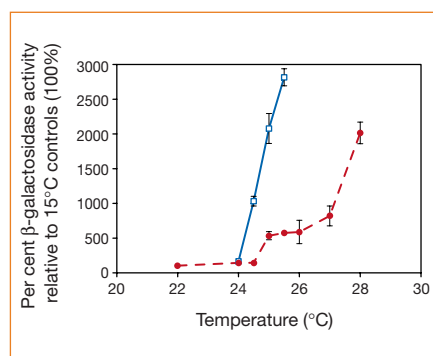


Figure 1 Saline⁹ suspensions of young adult PC72 worms grown synchronously at 15 °C (ref. 10) were split between three conditions for a total of 18 h: (1) exposed to microwaves (in TEM cell at 750 MHz and 0.5 W; ref. 5) within a Leec LT3 incubator; (2) temperature controls shielded with aluminium foil in the same incubator; (3) baseline controls at 15 °C. Incubator temperatures of 24.0, 24.5, 25.0 and 25.5 °C were tested using 12 replicates for each condition; controls only (6 replicates of condition 2) were also run at 22, 26, 27 and 28 °C. All worm samples were assayed fluorometrically^{4,5} for β-galactosidase activity. Enzyme activities were normalized against 15 °C baseline controls (100%) within each batch to allow comparison of reporter induction at different temperatures. Squares, blue solid line; reporter activities (± s.e.m.) in microwave-exposed cultures. Circles, red dashed line; control reporter activities (± s.e.m.) at each temperature.

We isolated the rotating oligomer from spinach-chloroplast ATP synthase enzyme preparations and reconstituted it at a high protein-to-lipid ratio in two-dimensional arrays. During the detergent-mediated procedure, the membrane-embedded oligomer remains intact (relative molecular mass 100K; individual subunits, known as subunits III, are 8.0K, and are equivalent to subunit-c in mitochondria and bacteria). The flanking stator proteins I, II and IV of this turbine, as well as all subunits ($\alpha_3\beta_3\gamma\delta\epsilon$) of the ATP-generating hydrophilic portion of the enzyme (known as F_1), are absent (Fig. 1a).

Imaging the topography of the densely packed subunit-III_x oligomers by contact-mode atomic force microscopy reveals alternating ring-like structures of two different diameters (Fig. 2). The narrow and the wider rings of the subunit-III oligomers have outer diameters of 5.9 ± 0.3 nm and 7.4 ± 0.3 nm, respectively. Both orifices have inner diameters of 3.5 ± 0.3 nm. Only a single, sharp band, corresponding to a relative molecular mass of 100K, is resolved on SDS-polyacrylamide gels of the protein arrays, which is the size expected for the subunit-III_x oligomer of the native ATP synthase from chloroplasts (Fig. 1a), indicating that most of the oligomers must have the same stoichiometry.

We conclude that the adjacent wide and narrow rings represent the two channel entrances of III_x oligomers, with opposite orientations perpendicular to the membrane surface (Fig. 2). The individual subunits of the 7.4-nm-diameter orifice protrude by 1.7 ± 0.3 nm from the bilayer surface, whereas those of the 5.9-nm-diameter orifice protrude by 1.5 ± 0.3 nm; the

oligomer, which is 7.3 ± 0.3 nm long, traverses the lipid bilayer (thickness, 4.1 ± 0.2 nm). Two III_x oligomers with wide orifices and two with narrow orifices are enlarged to display their substructure in more detail (Fig. 2). In most cases, 14 subunits per oligomer can be counted from the recorded images.

To determine the precise number of subunits in the circular III_x complex, we calculated angular power spectra from 320 individual images of well-preserved particles, which yielded a distinct single peak at 14-fold symmetry (Fig. 1c; filled circles). We then averaged data from the wide rings and the narrow rings (Fig. 1b) and their respective angular power spectra revealed the 14-fold symmetry. There was a weaker 12-fold signal (22% of the 14-fold power) still present, but no significant 7-fold to 13-fold symmetric contribution (Fig. 1c). This 14-fold symmetry and the same subunit-III oligomer dimensions were also evident in samples prepared by a different method (from Coomassie-blue non-denaturing gels without SDS), in which the stator protein IV was retained (data not shown).

We conclude that the proton turbine in the F_0 stem of chloroplast ATP synthase is an asymmetric cylindrical structure with 14 symmetrically distributed subunits that protrude from both membrane surfaces (Fig. 2).

So far the number of subunits, and therefore the diameter of the proton turbine, seems to be species-dependent^{5,6}. Subunit stoichiometry might even vary with metabolic state within the same organism, as indicated by biochemical studies on *E. coli*⁷. The proposed elasticity of the transmission between F_0 and F_1 could accommo-

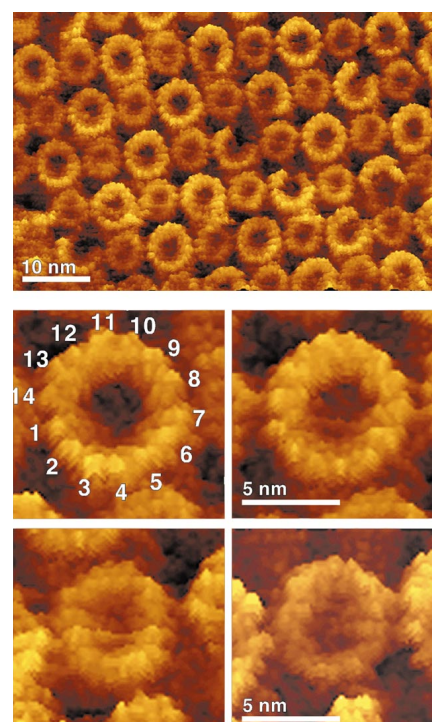


Figure 2 Subunit-III oligomers of chloroplast ATP synthase visualized in 25 mM MgCl₂, 10 mM Tris-HCl, pH 7.8, at room temperature using atomic force microscopy (Nanoscope III, Digital Instruments)¹¹. Top, the distinct wide and narrow rings represent the two surfaces of the subunit-III_x oligomer; middle, wide oligomer ends, showing 14 subunits-III; bottom, narrow oligomer ends. The full grey-level range of these topographs was 2 nm.

date such a variable stoichiometry⁸, as well as a non-integral H^+ /ATP ratio caused by symmetry mismatch between the three-fold symmetry of catalytic sites in F_1 and the 14-fold (10-fold⁶) symmetry of the F_0 oligomer.

Holger Seelert*, Ansgar Poetsch*, Norbert A. Dencher*, Andreas Engel†, Henning Stahlberg†, Daniel J. Müller†‡

*Physical Biochemistry, Department of Chemistry, Darmstadt University of Technology, D-64287 Darmstadt, Germany

e-mail: dencher@pop.tu-darmstadt.de

†M. E. Müller-Institute for Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

‡Max-Planck-Institute of Molecular Cell Biology and Genetics, D-01307 Dresden, Germany

1. Sambongi, Y. *et al.* *Science* **286**, 1722–1724 (1999).
2. Junge, W. *Proc. Natl Acad. Sci. USA* **96**, 4735–4737 (1999).
3. Dimroth, P., Wang, H., Grabe, M. & Oster, G. *Proc. Natl Acad. Sci. USA* **96**, 4924–4929 (1999).
4. Rastogi, V. K. & Girvin, M. E. *Nature* **402**, 263–268 (1999).
5. Jones, P. C. & Fillingame, R. H. *J. Biol. Chem.* **273**, 29701–29705 (1998).
6. Stock, D., Leslie, A. G. W. & Walker, J. W. *Science* **286**, 1700–1705 (1999).
7. Schemdt, R. A., Qu, J., Williams, J. R. & Brusilow, W. S. A. *J. Bacteriol.* **180**, 3205–3208 (1998).
8. Cherepanov, D. A., Mulkidjanian, A. Y. & Junge, W. *FEBS Lett.* **449**, 1–6 (1999).
9. Fromme, P., Boekema, E. J. & Gräber, P. Z. *Naturforsch.* **42C**, 1239–1245 (1987).
10. Neff, D. *et al.* *J. Struct. Biol.* **119**, 139–148 (1997).
11. Müller, D. J., Fotiadis, D., Scheuring, S., Müller, S. A. & Engel, A. *Biophys. J.* **76**, 1101–1111 (1999).

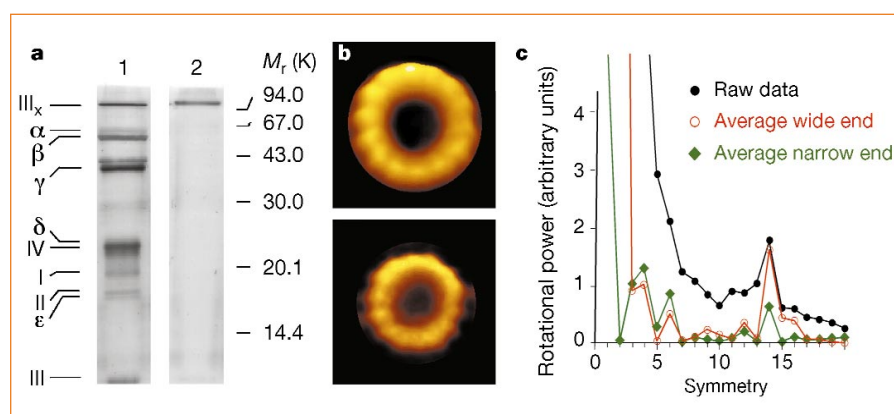


Figure 1 Characterization of subunit-III oligomers. **a**, SDS-PAGE and silver staining of chloroplast ATP synthase samples before and after reconstitution. Lane 1, the ATP synthase: proteins of the F_1 complex are designated by Greek letters, those of the F_0 complex by roman numerals. The sample was incubated in SDS buffer at 50 °C for 10 min to dissociate part of the III_x oligomer into its 8K monomers for reference. Lane 2, in reconstituted protein arrays, as used for atomic force microscopy, only the 100K III_x oligomer is present. To isolate intact subunit-III oligomers, SDS detergent⁹ was added to the CF₃F₂ complex obtained from rate-zonal centrifugation¹⁰ and then replaced with dodecyl maltoside detergent. The III_x oligomer was reconstituted into lipid bilayers (phosphatidylcholine and phosphatidic acid from egg yolk) by removing detergent with BioBeads SM2 (Bio-Rad, Hercules, California). **b**, Averaged topographs of the wide (top; $n = 220$) and narrow (bottom; $n = 220$) oligomer ends, generated by reference-free translational and rotational alignment of the individual particles. **c**, Rotational power spectra calculated from 320 topographs of individual oligomers (filled circles) and of averaged wide (circles) and narrow (diamonds) rings. The four-fold contribution of the spectra is due to the four neighbouring complexes in the crystal.

Roles of PPARs in health and disease

Sander Kersten, Béatrice Desvergne & Walter Wahli

Institut de Biologie Animale, Bâtiment de Biologie, Université de Lausanne, CH-1015 Lausanne, Switzerland

In developed societies, chronic diseases such as diabetes, obesity, atherosclerosis and cancer are responsible for most deaths. These ailments have complex causes involving genetic, environmental and nutritional factors. There is evidence that a group of closely related nuclear receptors, called peroxisome proliferator-activated receptors (PPARs), may be involved in these diseases. This, together with the fact that PPAR activity can be modulated by drugs such as thiazolidinediones and fibrates, has instigated a huge research effort into PPARs¹. Here we present the latest developments in the PPAR field, with particular emphasis on the physiological function of PPARs during various nutritional states, and the possible role of PPARs in several chronic diseases.

The PPARs were first cloned as the nuclear receptors that mediate the effects of synthetic compounds called peroxisome proliferators on gene transcription. It soon became clear that eicosanoids and fatty acids can also regulate gene transcription through PPARs. At the molecular level, PPARs act in a similar manner to other nuclear hormone receptors. First, they bind a specific element in the promoter region of target genes. PPAR and some other nuclear hormone receptors bind the promoter only as a heterodimer with the receptor for 9-*cis* retinoic acid, RXR (retinoid X receptor). Second, they activate transcription in response to binding of the hormone (ligand) (Fig. 1a). For the PPAR:RXR heterodimer, binding of the ligand of either receptor can activate the complex, but binding of both ligands simultaneously is more potent¹.

Three PPAR isotypes have been identified: α , β (also called δ and NUC1) and γ . PPAR α is expressed most in brown adipose tissue and liver, then kidney, heart and skeletal muscle. PPAR γ is mainly expressed in adipose tissue, and to a lesser extent in colon, the immune system and the retina. PPAR β is found in many tissues but the highest expression is in the gut, kidney and heart¹.

PPARs are ligand-dependent transcription factors: activation of target gene transcription depends on the binding of the ligand to the receptor. Some ligands are shared by the three isotypes, such as polyunsaturated fatty acids and probably oxidized fatty acids. Several compounds bind with high affinity to PPAR α , including long-chain unsaturated fatty acids such as linoleic acid, branched, conjugated and oxidized fatty acids such as phytanic acid and conjugated linolenic acid, and eicosanoids such as 8S-HETE and leukotriene (LT) B₄¹⁻⁴ (Fig. 1b). This last compound is particularly interesting, as a membrane receptor for LTB₄ has also been cloned. The functional relationship between these two types of receptor is not clear. The prostaglandin 15-deoxy-D^{12,14}-prostaglandin J₂ is the most potent natural ligand of PPAR γ , but the extent to which its *in vivo* effects are mediated through PPAR γ is not known.

For a review of PPARs, particularly their molecular mode of action, see ref. 1 and references therein (including many not cited here for reasons of space).

PPAR function at the cellular level

Much of the function of PPARs can be extrapolated from the identity of their target genes, which so far all belong to pathways of lipid transport and metabolism. PPAR α has mostly been studied in the context of liver parenchymal cells, where it is highly expressed. The target genes of PPAR α are a relatively homogenous group of genes that participate in aspects of lipid catabolism such as fatty acid uptake through membranes, fatty acid binding in cells, fatty acid oxidation (in microsomes, peroxisomes and mitochondria) and lipoprotein assembly and transport (Fig. 2).

Whereas PPAR α operates in the catabolism of fatty acids in the liver, PPAR γ influences the storage of fatty acids in the adipose tissue (Fig. 2). With the C/EBP transcription factors, PPAR γ is part of the adipocyte differentiation program that induces the maturation

of pre-adipocytes into fat cells⁵. Most of the PPAR γ target genes in adipose tissue are directly implicated in lipogenic pathways, including lipoprotein lipase (LPL), adipocyte fatty acid binding protein (A-FABP or aP2), acyl-CoA synthase and fatty acid transport protein (FATP).

PPAR β has received little attention, probably because of the lack of a connection with important clinical manifestations. However, PPAR β is linked to colon cancer (see below), among other functions. PPAR β regulates the expression of acyl-CoA synthetase 2 in the brain, linking PPAR β to basic lipid metabolism⁶. Moreover, it probably participates in embryo implantation and decidualization⁷. These data will spur new interest in the study of PPAR β function.

PPARs in whole body physiology

The role of these transcription factors in whole body human physiology and metabolism can best be illustrated by comparing two opposite nutritional states: early absorptive period or fed state and late post-absorptive period or fasting state. In the fed state, which in humans is up to 4 h after a large meal, carbohydrates and fat enter the circulation in the form of glucose and chylomicrons, respectively (Fig. 3a). Most glucose is taken up by the liver and, if

Box 1

Gene targeting and human molecular genetics

Use of genetic approaches such as gene targeting and human molecular genetics has led to major advances in our knowledge about the physiological role and medical significance of PPAR γ . Unlike PPAR α null mice, which have no obvious phenotype unless fasted, null mice for PPAR γ die as embryos^{13,26}, PPAR γ null embryos that are rescued to term have no visible white adipose tissue and a fatty liver²⁶. Also, embryonic stem cells lacking PPAR γ are unable to differentiate into adipocytes *in vitro*⁵. Both results show the contribution of PPAR γ to adipogenesis.

The function of PPAR γ in improving insulin sensitivity is controversial. In disagreement with the purported anti-diabetic effect of PPAR γ , heterozygous PPAR γ mice, which are viable, are less prone to insulin resistance on a high fat diet¹³. Remarkably, this protection is lost when these animals are treated with pioglitazone, a synthetic PPAR γ ligand belonging to the TZD class. Tissue-specific gene targeting should soon clarify some of the discrepancies.

Human genetic studies give important clues about the function of PPAR γ in mammalian metabolism and its link to certain chronic diseases. Four single amino-acid substitutions within human PPAR γ have been described. A P12A mutation at the extreme amino-terminus of PPAR γ is most common, but its effect on weight gain and insulin sensitivity is unclear²⁷⁻²⁹. A P115Q mutation was identified in four extremely obese patients with surprisingly little defect in insulin sensitivity. Very recently, two new mutations were found in three patients suffering from severe insulin resistance but not obesity³⁰. Overall the data suggest that PPAR γ promotes fat storage and insulin sensitivity, but additional studies are needed to confirm these findings.

glycogen stores are already filled, used for lipogenesis. The amount of the transcription factor sterol response element binding protein 1 (SREBP1, also called ADD1) rises in the fed state, which promotes the glycolytic conversion of glucose into acetyl-CoA and subsequently the synthesis of fatty acids from acetyl-CoA^{8,9}. Fatty acids are converted to triglycerides and packaged into very low density lipoproteins (VLDL).

In adipose tissue, the amounts of SREBP and PPAR γ are elevated, probably because of regulation by insulin¹⁰. PPAR γ is a direct target gene of SREBP¹¹, which emphasizes the cooperative and additive functions between these two types of receptor. In addition, SREBP1 may be involved in producing an endogenous ligand (probably fatty acid) for PPAR γ . The overall effect is stimulation of the uptake of glucose and fatty acids, and their subsequent conversion to triglycerides.

Triglyceride storage causes increased production of the hormone leptin in the adipose tissue. Leptin is the protein product of the *ob* gene, whose deletion leads to severe obesity in mice. Its expression is increased by long-term overfeeding as part of a feedback mechanism to limit further food intake and weight gain. Consistent with the role of PPAR γ in promoting lipogenesis, production of leptin in adipose tissue is under negative control by PPAR γ . Paradoxically, expression of both PPAR γ and leptin is reduced by fasting and increased by feeding. In the latter case, PPAR γ may attenuate the increase in

leptin expression to limit wasteful lipolysis and fatty acid oxidation, processes which are stimulated by leptin¹². If the above scenario is correct, decreased PPAR γ expression may lead to increased leptin levels and, as a result, to lower food intake and weight gain. Studies with PPAR γ ^{+/-} mice indicate that this is the case¹³.

A different situation exists in the late post-absorptive or fasting state (Fig. 3b). In the liver, fatty acids are oxidized to acetyl-CoA and subsequently to ketone bodies, such as acetoacetate and β -hydroxybutyrate. Both processes are strongly stimulated by PPAR α , expression of which is elevated upon fasting¹⁴. Fatty acids are ligands for PPARs, so it is possible that the large amounts of fatty acids liberated from the adipose tissue can stimulate their own metabolism by activating PPAR α . Experiments with PPAR α null mice show that PPAR α is important in the hepatic response to fasting. When fasting, these mice suffer from a defect in fatty acid oxidation and ketogenesis, resulting in elevated plasma free fatty acids, hypoketonaemia, hypothermia and hypoglycaemia^{14,15}. The hypoglycaemia emphasizes the important interplay between fatty acid and glucose metabolism in energy homeostasis.

In adipose tissue, expression of SREBP and PPAR γ is low under fasting conditions. Under a strong adrenergic stimulus, triglycerides are hydrolysed to fatty acids and glycerol, but some of the released fatty acids are re-esterified to triglycerides, in a reaction that requires synthesis of glycerol from gluconeogenic precursors. The rate-limiting step for glyceroneogenesis is catalysed by phosphoenolpyruvate carboxykinase, whose transcription is positively controlled by PPAR γ . Thus, even under highly catabolic conditions such as fasting, lipogenesis continues and is dependent upon PPAR γ .

Therapeutic potential of PPAR ligands

In developed societies, metabolic disorders such as hyperlipidaemia, atherosclerosis, diabetes and obesity rarely occur in isolation, but are usually part of a complex phenotype of metabolic abnormalities called syndrome X. Synthetic agonists for both PPAR α (fibrates) and PPAR γ (thiazolidinediones; TZDs) are useful in the treatment of the diseases that are part of this syndrome.

Synthetic PPAR γ ligands are used for their potent antidiabetic effects. In the United States, three TZDs, troglitazone (Rezulin), rosiglitazone (Avandia) and pioglitazone (Actos), are approved for use in type II diabetic patients. They bind PPAR γ with moderate (troglitazone) to high (rosiglitazone) affinity, so it is believed that their hypoglycaemic effect is exerted by activating PPAR γ . However,

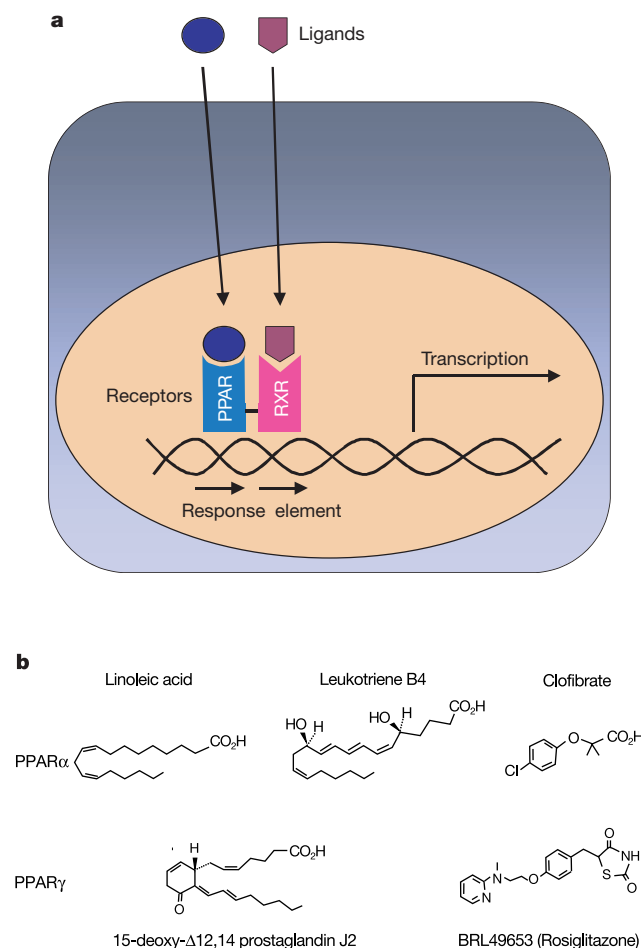


Figure 1 PPARs at the basic level. **a**, Basic mechanism of action of nuclear hormone receptors. Nuclear hormone receptors bind to a specific sequence in the promoter of target genes (called hormone response elements), and activate transcription upon binding of ligand. Several nuclear hormone receptors, including the retinoic acid receptor, the vitamin D receptor and PPAR, can bind to DNA only as a heterodimer with the retinoid X receptor, RXR, as shown. **b**, Structure of some PPAR α and PPAR γ ligands.

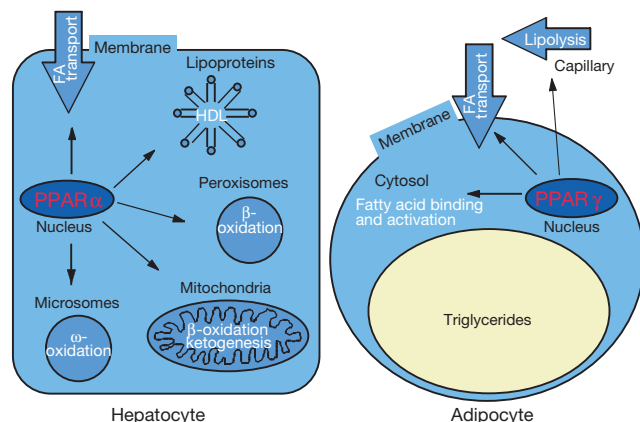


Figure 2 Action of PPAR α and PPAR γ at the cellular level. PPAR α stimulates oxidation of fatty acids in various organelles, such as mitochondria, peroxisomes and microsomes. It also stimulates uptake of fatty acids and synthesis of lipoproteins. PPAR γ stimulates lipolysis of circulating triglycerides and the subsequent uptake of fatty acids into the adipose cell. It also stimulates binding and activation of fatty acids in the cytosol, events that are required for synthesis of triglycerides. FA, fatty acid; HDL, high density lipoprotein.

a direct connection between PPAR γ and glucose homeostasis has not been easy to establish because skeletal muscle, which accounts for the TZD-mediated increase in glucose disposal, expresses only trace amounts of PPAR γ . To explain this paradox, a mechanism has been proposed by which TZDs divert fatty acids away from skeletal muscle by increasing their uptake in adipose tissue, and so reduce the deleterious effects of fatty acids on muscle insulin action. In this model, the hypoglycaemic effect of TZDs is secondary to their hypolipidaemic effect. However, mice that lack adipose tissue can still benefit from the action of TZDs, indicating that adipose tissue is dispensable for mediating the hypoglycaemic effects of TZDs. It is also possible that the effect of TZDs on glucose homeostasis might be via an alternative mechanism not involving PPAR γ , as is the case for the inhibitory effect of troglitazone on cholesterol synthesis¹⁶. New ligands specific for PPAR γ will be useful in refining these observations and working out the mechanisms of glucose homeostasis.

Fibrates are potent hypolipidaemic drugs. In the past few years they have been used increasingly to treat cardiovascular disease. Fibrates, which include gemfibrozil, bezafibrate and fenofibrate, bind PPAR α with high affinity and it is believed that most of their effects on disease progression are mediated by PPAR α . Fibrates lower plasma triglyceride levels markedly and increase high density

lipoprotein (HDL) levels. The former effect occurs by stimulating hepatic fatty acid oxidation and reducing apoCIII expression, whereas the latter effect is due to induction of apolipoprotein-AI and apolipoprotein-AII expression, both mediated by PPAR α . Fibrates may also have a hypoglycaemic and thus anti-diabetic effect, as a consequence of their hypolipidaemic action. Accordingly, it is interesting to compare PPAR α agonists and PPAR γ agonists. Though they act through different receptor isotypes, both groups of compound have potent hypolipidaemic properties which may be at the basis of their hypoglycaemic effect. Future studies will have to establish whether fibrates or other PPAR α agonists may be applicable in the treatment of type II diabetes.

The heterodimer between PPAR and RXR can also be activated by ligand binding to RXR. Synthetic ligands specific for RXR have been tested as an alternative treatment for diabetes. Compounds such as LG1069 and LG100268 have a potent hypoglycaemic effect in animal studies. The effects of these RXR agonists on glucose homeostasis have not been fully established, but they are probably mediated by the PPAR:RXR heterodimer.

Atherosclerosis

Atherosclerosis is a complex disease to which many factors contribute. It is characterized by a gradual build-up of lipids in the arterial wall (atherosclerotic plaque), often leading to sudden obstruction of blood flow after rupture of this plaque. Endothelial dysfunction resulting in chronic inflammation of the vascular wall, proliferation of smooth muscle cells and formation of foam cells is important in the development of atherosclerosis. As well as regulating plasma lipoprotein concentrations, PPAR α and PPAR γ may affect foam cell formation, modulate the inflammatory response and influence plaque stability. PPAR α may also decrease the plasma concentration of pro-atherosclerotic proteins such as fibrinogen and C-reactive protein¹⁷.

In 1996 it was proposed that PPAR α may be involved in inflammation, as it is the nuclear receptor for the eicosanoid LTB₄. Mice lacking PPAR α display a prolonged response to an inflammatory stimulus, indicating that PPAR α has an anti-inflammatory action¹⁸. Fibrates can modulate inflammation by inhibiting cytokine (TNF α , interleukins) production in a PPAR α -dependent manner¹⁹. In contrast, fibrates markedly increase plasma tumour necrosis factor- α (TNF α) levels in mice, an effect also mediated by PPAR α ²⁰. However, in rodents increased TNF α may be secondary to the hypolipidaemic and peroxisome proliferative effects of fibrates. It is clear that PPAR α plays a role in the inflammatory process but much still needs to be learned about the details of its function.

PPAR γ has also been implicated in inflammation. In monocytes/macrophages, PPAR γ has been proposed to reduce cytokine (TNF α , interleukin-1 β , interleukin-6) production by inhibiting the activity of pro-inflammatory transcription factors such as AP-1, STAT and NF- κ . This anti-inflammatory effect of PPAR γ could be beneficial in the treatment of atherosclerosis. In addition, PPAR γ may reduce expression of metalloproteinases such as MMP-9, which are implicated in plaque destabilization²¹.

In contrast, PPAR γ may promote atherosclerosis by stimulating the uptake of oxidized LDL, a critical event in foam cell formation. A feed-forward mechanism has been proposed whereby oxidized fatty acids that enter the cell through oxidized LDL can activate PPAR γ , further stimulating uptake of oxidized LDL.

However, the effect of TZDs and 15-deoxy-D^{12,14}-prostaglandin J₂ on atherosclerosis and inflammation do not always correspond²². This suggests that at least one type of ligand, probably 15-deoxy-D^{12,14}-prostaglandin J₂, acts independently of PPAR γ , although PPAR γ -independent effects of TZDs have been reported as well¹⁶. Use of antisense and/or gene-targeting technology should clarify the role of PPAR γ in inflammation and atherosclerosis. Without this information and without clear data from clinical trials, the link between PPAR γ and atherosclerosis remains speculative.

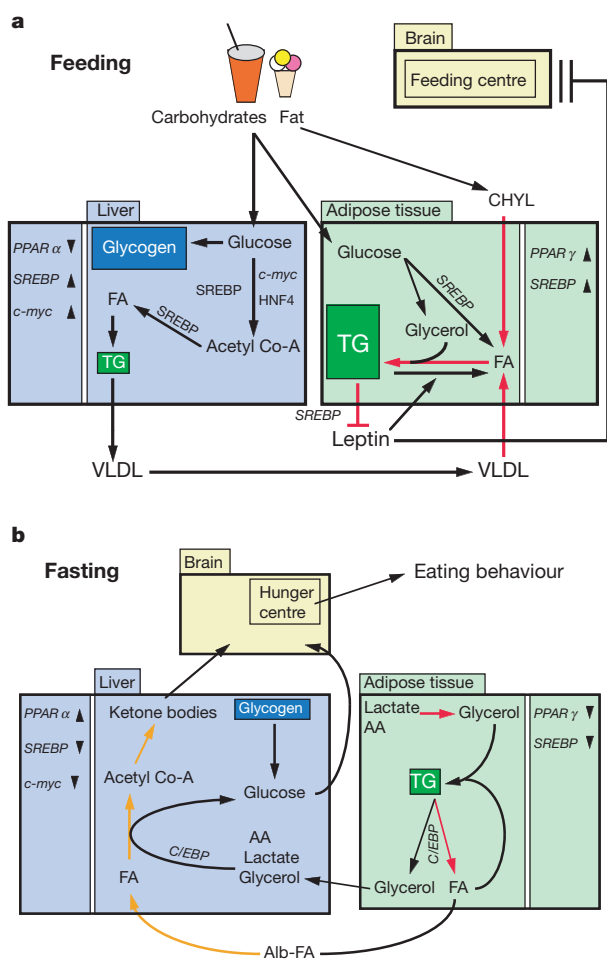


Figure 3 Overview of the roles of PPAR α and PPAR γ during feeding and fasting. **a**, PPAR α is mainly active in the fed state when it stimulates accumulation of fat in the adipose tissue. **b**, PPAR α is mainly active in the fasting state when it stimulates oxidation of fatty acids in the liver. The particular metabolic steps that are regulated by PPAR α and PPAR γ are shown in orange and red, respectively. The roles of the transcription factors C/EBP, c-myc, HNF4 and SREBP, which are transcriptional regulators of metabolism, have been illustrated. The expression of the indicated proteins is increased ($\blacktriangle$) or decreased ($\blacktriangledown$). AA, amino acids; CHYL, chylomicrons; FA, fatty acids; TG, triglycerides.

Cancer

PPAR α mediates the hepatocarcinogenic effect of certain peroxisome proliferators in rodents. However, no carcinogenic effect of peroxisome proliferators has been found in humans, possibly because expression of PPAR α is much lower in human liver than in rodent liver or due to other species-specific differences.

PPAR γ has an anti-proliferative effect in pre-adipocytes and possibly in several malignant cell types. PPAR γ ligands can induce terminal differentiation of human liposarcoma cells *in vitro* and in patients suffering from advanced liposarcoma²³. PPAR γ ligands also promote terminal differentiation of malignant breast epithelial cells *in vitro*, induce apoptosis and fibrosis of injected breast tumour cells (MCF-7) in mice, and reduce tumour incidence in rats treated with nitrosomethylurea²⁴. An anti-tumour effect of PPAR γ ligands was also observed in mice injected with prostate tumour cells (PC-3). Unfortunately, the picture is less clear for colon cancer. PPAR γ ligands have been reported both to promote and to protect against colon cancer in mice. Interestingly, PPAR β has also been linked to colon cancer. It is a negative target of the APC gene, which is mutated in familial adenomatous polyposis, an inherited disease characterized by numerous colorectal adenomas²⁵. In addition, the nonsteroidal anti-inflammatory drug (NSAID) sulindac, which suppresses colorectal tumorigenesis, can antagonize PPAR β . Thus PPAR β may be a critical intermediate in the tumorigenic pathway of the APC gene and be the molecular target for the effect of NSAIDs on colorectal cancer.

There are however two problems associated with the type of studies described above. First, it is difficult to extrapolate *in vitro* data and data from animal models to the human situation. Second, they make the possibly false assumption that all of the effects of PPAR ligands are mediated through activation of PPAR.

Conclusion

Since their discovery in the early 1990s it has become clear that PPARs are crucial in the genetic regulation of complex pathways of mammalian metabolism, including fatty acid oxidation and lipogenesis. Whereas PPAR α promotes fatty acid oxidation under conditions of lipid catabolism such as fasting, PPAR γ acts at the level of the adipose tissue and promotes lipogenesis under anabolic conditions. Much research is directed towards the identification of high-affinity, high-specificity agonists and antagonists for the treatment of hyperglycaemia, hyperlipidaemia and other metabolic diseases. Further advances in modern technology should assist in answering some of the more pertinent questions relating to PPARs, particularly with respect to the purported role of PPARs in atherosclerosis and cancer. □

- Desvergne, B. & Wahli, W. Peroxisome proliferator-activated receptors: nuclear control of metabolism. *Endocr. Rev.* **20**, 649–688 (1999).
- Lin, Q., Ruuska, S. E., Shaw, N. S., Dong, D. & Noy, N. Ligand selectivity of peroxisome proliferator activated receptor alpha. *Biochemistry* **38**, 185–190 (1999).
- Ellinghaus, P., Wolfrum, C., Assmann, G., Spencer, F. & Seedorf, U. Phytanic acid activates the peroxisome proliferator-activated receptor alpha (PPARalpha) in sterol carrier protein 2-sterol carrier protein x-deficient mice. *J. Biol. Chem.* **274**, 2766–2772 (1999).
- Moya-Camarena, S. Y., Vanden Heuvel, J. P., Blanchard, S. G., Leesnitzer, L. A. & Belury, M. A. Conjugated linoleic acid is a potent naturally occurring ligand and activator of PPARalpha. *J. Lipid Res.* **40**, 426–433 (1999).

- Rosen, E. D. *et al.* PPAR is required for the differentiation of adipose tissue *in vivo* and *in vitro*. *Mol. Cell* **4**, 611–617 (1999).
- Basu-Modak, S. *et al.* Peroxisome proliferator-activated receptor beta regulates acyl-CoA synthetase 2 in reaggregated rat brain cell cultures. *J. Biol. Chem.* **274**, 35881–35888 (1999).
- Lim, H. *et al.* Cyclo-oxygenase-2-derived prostacyclin mediates embryo implantation in the mouse via PPAR δ . *Genes Dev.* **13**, 1561–1574 (1999).
- Foretz, M. *et al.* ADD1/SREBP-1c is required in the activation of hepatic lipogenic gene expression by glucose. *Mol. Cell. Biol.* **19**, 3760–3768 (1999).
- Foretz, M., Guichard, C., Ferre, P. & Foufelle, F. Sterol regulatory element binding protein-1c is a major mediator of insulin action on the hepatic, expression of glucokinase and lipogenesis-related genes. *Proc. Natl Acad. Sci. USA* **96**, 12737–12742 (1999).
- Rieusset, J. *et al.* Insulin acutely regulates the expression of the peroxisome proliferator-activated receptor-gamma in human adipocytes. *Diabetes* **48**, 699–705 (1999).
- Fajas, L. *et al.* Regulation of peroxisome proliferator-activated receptor gamma expression by adipocyte differentiation and determination factor 1/sterol regulatory element binding protein 1: implications for adipocyte differentiation and metabolism. *Mol. Cell. Biol.* **19**, 5495–5503 (1999).
- Wang, M. Y., Lee, Y. & Unger, R. H. Novel form of lipolysis induced by leptin. *J. Biol. Chem.* **274**, 17541–17544 (1999).
- Kubota, N. *et al.* PPAR mediates high-fat diet-induced adipocyte hypertrophy and insulin resistance. *Mol. Cell* **4**, 597–609 (1999).
- Kersten, S. *et al.* Peroxisome proliferator activated receptor alpha mediates the adaptive response to fasting. *J. Clin. Invest.* **103**, 1489–1498 (1999).
- Leone, T. C., Weinheimer, C. J. & Kelly, D. P. A critical role for the peroxisome proliferator-activated receptor alpha (PPARalpha) in the cellular fasting response: the PPARalpha-null mouse as a model of fatty acid oxidation disorders. *Proc. Natl Acad. Sci. USA* **96**, 7473–7478 (1999).
- Wang, M., Wise, S. C., Leff, T. & Su, T. Z. Troglitazone, an antidiabetic agent, inhibits cholesterol biosynthesis through a mechanism independent of peroxisome proliferator-activated receptor-gamma. *Diabetes* **48**, 254–260 (1999).
- Kockx, M. F. *et al.* Fibrates suppress fibrinogen gene expression in rodents via activation of the peroxisome proliferator-activated receptor-alpha. *Blood* **93**, 2991–2998 (1999).
- Devchand, P. R. *et al.* Chemical probes that differentially modulate peroxisome proliferator-activated receptor alpha and BLTR, nuclear and cell surface receptors for leukotriene B(4). *J. Biol. Chem.* **274**, 23341–23348 (1999).
- Delaville, P. *et al.* Peroxisome proliferator-activated receptor alpha negatively regulates the vascular inflammatory gene response by negative cross-talk with transcription factors NF-kappaB and AP-1. *J. Biol. Chem.* **274**, 32048–32054 (1999).
- Hill, M. R. *et al.* Effect of peroxisome proliferator-activated receptor alpha activators on tumor necrosis factor expression in mice during endotoxemia. *Infect. Immun.* **67**, 3488–3493 (1999).
- Ricote, M., Huang, J. T., Welch, J. S. & Glass, C. K. The peroxisome proliferator-activated receptor (PPARgamma) as a regulator of monocyte/macrophage function. *J. Leukoc. Biol.* **66**, 733–739 (1999).
- Thieringer, R. *et al.* Activation of peroxisome proliferator-activated receptor gamma does not inhibit IL-6 or TNF- α responses of macrophages to lipopolysaccharide *in vitro* and *in vivo*. *J. Immunol.* **164**, 1046–1054 (2000).
- Demetri, G. D. *et al.* Induction of solid tumor differentiation by the peroxisome proliferator-activated receptor-gamma ligand troglitazone in patients with liposarcoma. *Proc. Natl Acad. Sci. USA* **96**, 3951–3956 (1999).
- Suh, N. *et al.* A new ligand for the peroxisome proliferator-activated receptor-gamma (PPAR-gamma), GW7845, inhibits rat mammary carcinogenesis. *Cancer Res.* **59**, 5671–5673 (1999).
- He, T. C., Chan, T. A., Vogelstein, B. & Kinzler, K. W. PPARdelta is an APC-regulated target of nonsteroidal anti-inflammatory drugs. *Cell* **99**, 335–345 (1999).
- Barak, Y. *et al.* PPAR gamma is required for placental, cardiac, and adipose tissue development. *Mol. Cell* **4**, 585–595 (1999).
- Valve, R. *et al.* Two polymorphisms in the peroxisome proliferator-activated receptor-gamma gene are associated with severe overweight among obese women. *J. Clin. Endocrinol. Metab.* **84**, 3708–3712 (1999).
- Ek, J. *et al.* Homozygosity of the Pro12A1a variant of the peroxisome proliferation-activated receptor-gamma2 (PPAR-gamma2): divergent modulating effects on body mass index in obese and lean Caucasian men. *Diabetologia* **42**, 892–895 (1999).
- Mancini, F. P. *et al.* Pro12A1a substitution in the peroxisome proliferator-activated receptor-gamma2 is not associated with type 2 diabetes. *Diabetes* **48**, 1466–1468 (1999).
- Barroso, I. *et al.* Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* **402**, 880–883 (1999).

Acknowledgements

S.K. was supported by fellowships from the European Molecular Biology Organization and the Roche Research Foundation. Work in the author's laboratory is financed by the Swiss National Science Foundation, the Etat de Vaud, and Human Frontier Science Programme.

Correspondence should be addressed to W.W. (e-mail: walter.wahli@iba.unil.ch).

Neoproterozoic 'snowball Earth' simulations with a coupled climate/ice-sheet model

William T. Hyde*, Thomas J. Crowley*, Steven K. Baum* & W. Richard Peltier†

* Department of Oceanography, Texas A&M University, College Station, Texas 77843-3146, USA

† Department of Physics, University of Toronto, Toronto, Ontario, M5S 1A7, Canada

Ice sheets may have reached the Equator in the late Proterozoic era (600–800 Myr ago), according to geological and palaeomagnetic studies, possibly resulting in a 'snowball Earth'. But this period was a critical time in the evolution of multicellular animals, posing the question of how early life survived under such environmental stress. Here we present computer simulations of this unusual climate stage with a coupled climate/ice-sheet model. To simulate a snowball Earth, we use only a reduction in the solar constant compared to present-day conditions and we keep atmospheric CO₂ concentrations near present levels. We find rapid transitions into and out of full glaciation that are consistent with the geological evidence. When we combine these results with a general circulation model, some of the simulations result in an equatorial belt of open water that may have provided a refugium for multicellular animals.

Some of the most dramatic events in the Earth's history occurred at the end of the Proterozoic era (this era was about 1,000–540 Myr ago). This epoch was characterized by formation of the supercontinent of Rodinia from about 1,000 to 800 Myr ago, the later breakup of this landmass and eventual reassembly into a different configuration by ~550 Myr ago, during the final phase of the pan-African orogeny^{1,2}. There were also major changes in strontium, sulphur and carbon isotopes³, together with the most extensive glaciation of the past billion years. At least two main phases of ice advance occurred, with glaciers apparently extending to the Equator at sea level^{4–8}. The first phase from ~760 to 700 Myr is generally termed the Sturtian ice age, while the second, which occurred in the interval ~620–580 Myr ago, is often termed the Varanger (or Marinoan) ice age.

The late Proterozoic also marked the first appearance of metazoans (multicelled animals), perhaps as early as 700–1,000 Myr ago^{9–12}, while the Varanger ice age was almost immediately followed by the time interval (Vendian) featuring the first fossil remains of multicelled animals—the Ediacaran fauna. If metazoans are pre-Varanger, extensive glaciation may have exerted a significant stress on biota during a critical interval in their evolution.

Recent work⁴ has focused attention on the Neoproterozoic by interpreting new carbon-isotope data to indicate that biological productivity of the oceans virtually ceased for perhaps millions of years during the glacial era. These authors concluded from this and other evidence that the planet entered a snowball Earth state, in which it was completely covered by ice until CO₂ outgassing produced a sufficiently large greenhouse effect to melt the ice. In this scenario the sudden warming caused a rapid precipitation of calcium carbonate, producing the cap carbonate rocks often observed in strata of this era. The repetition of such formations suggests that this sequence of events occurred at least twice in the Neoproterozoic.

Although this provocative hypothesis has stimulated considerable interest, it remains to be shown that the physical climate system can respond in the proposed manner. For example, although the concepts of the snowball Earth 'attractor', and rapid exits therefrom, are a feature of energy balance models (EBMs)¹³, it is not at all clear to what degree the slow dynamics of ice-sheet physics would modify this model. Previous simulations with EBMs¹⁴ and general circulation models (GCMs)^{15,16} indicated that plausible changes in solar luminosity, the Earth's rotation rate and ocean heat transport were

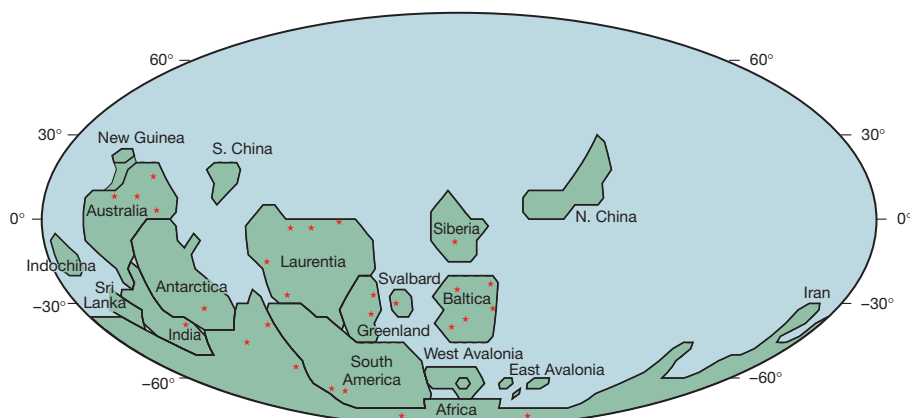


Figure 1 Late Precambrian geography. The figure shows the input geography² (Mollweide projection) used in our late Precambrian simulation. Red asterisks indicate the location of Neoproterozoic glacial deposits⁷.

insufficient to produce a snowball Earth unless CO_2 is lowered to less than half the present value. Simulations with a radically increased planetary obliquity—which has been suggested as an explanation for equatorial glaciation¹⁷—preferentially produce snow in low but not high latitudes¹⁸, and their seasonally extremely warm polar areas are in direct conflict with evidence for glaciation in these regions (see, for example, Fig. 1).

None of the above simulations included an ice-sheet model, and few evaluated the effects of orbital insolation variations or used 'realistic' geography. Here we demonstrate that a coupled EBM/ice-sheet model using realistic boundary conditions can successfully predict land ice at the Equator. We further demonstrate that a stand-alone general circulation model experiment with ice-covered continents predicts open water in equatorial regions, thereby creating a refugium for metazoans in this extreme climate.

Ice-sheet model and boundary conditions

The climate/ice-sheet model that we use was introduced by Deblonde and Peltier^{19,20}; here, we will use the version described fully by Tarasov and Peltier²¹. Its physical basis consists of four submodels, which predict ice flow, mass balance, temperature and bedrock sinking. In brief, ice is assumed to flow subject to a temperature-independent rheology based on the Nye²² formulation. Precipitation/ablation is computed according to statistical models^{23,24} which themselves take as input monthly temperatures from a nonlinear, two-dimensional (latitude–longitude) diffusive seasonal EBM²⁵. Bedrock sinking is assumed to occur with a time constant of 4,000 years. The EBM has been validated against many different GCMs²⁶, while the coupled EBM/ice-sheet model reproduces many features of climate change over the past 120,000 years and predicts ice for about 80% of the known glacial deposits for the radically different geography of the Carboniferous ice age²⁷. The flexibility of the model lends itself to extensive sensitivity testing and long runs; for example, here we have run approximately 20 million years of model simulations.

Model inputs include palaeogeography, atmosphere CO_2 concentrations, precipitation, Milankovitch forcing, and the solar constant. Our baseline palaeogeography for the ~590-Myr Varanger glaciation is taken from Dalziel's reconstruction for 545 Myr ago². Although data suggest that at 590 Myr ago amalgamation was not as great as it was at 545 Myr, the latitudes of important glacial sites in Namibia and Australia are close to their Varanger positions. However, interior seaways existed before the final joining of Africa and South America, and Baltica was closer to Laurentia before the opening of the Iapetus Ocean. We will later discuss a number of experiments that examine the sensitivity of our conclusions to

uncertainties in palaeogeography. For convenience we initially assume that the continents are at sea level; adding freeboard is nearly equivalent to reducing the solar constant. We employ a solar constant appropriate to the era in question, that is, about 6% below present¹⁴. Precipitation and CO_2 values are varied to test the snowball Earth hypothesis. In our initial simulations, ice sheets were driven by a uniform precipitation of 0.6 mm d^{-1} , which declines with height and decreasing temperature. This value is characteristic of current Northern Hemisphere mid-latitude land areas. Given the important role which the tropics play in our simulations, a lower value cannot be justified for our (initially ice-free) Neoproterozoic simulations. The model is relatively insensitive to higher precipitation rates²⁷. The model is driven with Milankovitch variations appropriate to the Pleistocene epoch²⁸, although the very cold climates we investigate are relatively uninfluenced by this forcing. The model makes predictions of temperature, sea ice, and ice volume that can be tested for consistency against the snowball Earth hypothesis and other geological evidence.

Neoproterozoic simulations

Our first experiment, with the above inputs and present-day atmospheric CO_2 concentration resulted in a large ice sheet (not shown) that in coastal areas approaches 40° latitude, but which is restricted by summer warming on the supercontinent to about 50°S in the interior. Sensitivity experiments with the alternative reconstruction of Bond *et al.*²⁹ yield similar results, but the smaller amount of land area in middle to high latitudes leads to ice growth to 30° palaeolatitude (not shown).

Neoproterozoic CO_2 levels are not known, but may have been very variable due to large carbon reservoir shifts as indicated by $\delta^{13}\text{C}$ changes^{3,4,30}. We therefore conducted a series of sensitivity tests in which CO_2 levels were varied by adding or subtracting infrared forcing. For CO_2 levels approximately half that of the Holocene epoch (that is, ~130 p.p.m. or a cooling of roughly 5 W m^{-2}) the entire land surface is glaciated (Fig. 2), while for higher levels (slightly over twice the Holocene value) the ice volume approaches that of the Last Glacial Maximum. The transition from a cold state to a fully snow-covered one (snowball Earth) is abrupt (~2,000 years; Fig. 3), as conjectured by Hoffman *et al.*⁴ and consistent with other EBM³¹, GCM³² and EBM/ice-sheet studies²⁷. It occurs at a CO_2 level equivalent to an infrared cooling of 4.75 to 5.0 W m^{-2} . The exact location of this discontinuity is not particularly important, given the known uncertainty in solar constant and—for example—the different land albedo of a world without vascular plants. The

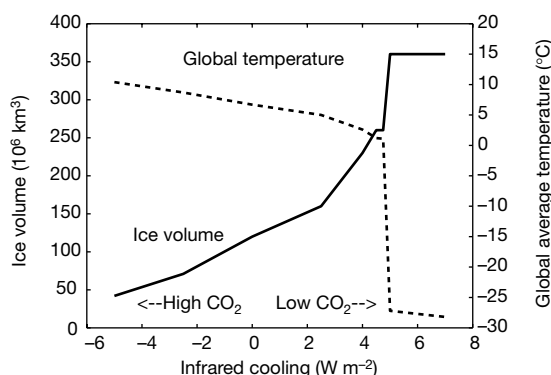


Figure 2 Operating curve for the climate/ice-sheet model. Ice volume and global sea level temperature are shown as a function of CO_2 level and of infrared forcing. A bifurcation occurs at a cooling of $\sim 5 \text{ W m}^{-2}$, beyond which all land surfaces are ice-covered and global temperature drops considerably: the snowball Earth.

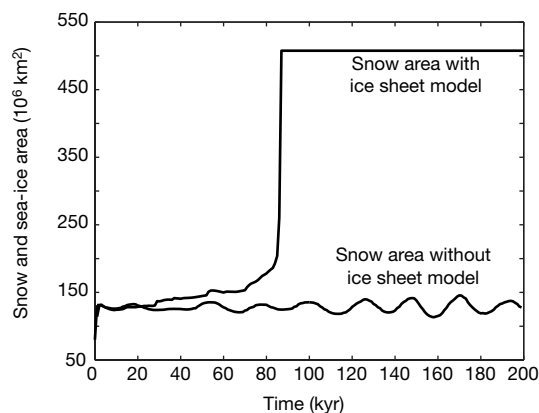


Figure 3 The effect of ice dynamics on model results for an experiment with -5 W m^{-2} of infrared forcing. Without the ice sheet, the climate model predicts a climate in which only about one-quarter of the planetary surface is covered in snow or sea ice. If the ice sheets are allowed to grow, however, they flow into areas originally too warm for year-round snow, but the ice-sheet's thermal inertia and elevation prevent it from melting.

transition is associated with an abrupt drop in globally averaged surface temperature, from approximately zero to about -27°C .

Continental freeboard and atmosphere CO_2 concentrations (or luminosity) play complementary roles in determining ice volume—the higher the freeboard, the greater the level of CO_2 consistent with ice-covered continents. Sensitivity experiments (not shown), in which we raised baseline continental topography by 100-m increments, indicate that for a 500-m baseline, the snowball Earth solution occurs with CO_2 levels approximately 50–100% greater than present. These are reasonable freeboards for the interval (800–550 Myr ago) associated with the pan-African orogeny.

A number of additional experiments (not shown) were run to determine the sensitivity of our results to continental configuration. For example, most of the Neoproterozoic glaciation occurred in an era of dispersed geography⁴. In the absence of a reliable 590-Myr geography, we tested the effect of continental dispersion by using a Cenomanian (94 Myr ago) geography³³. No qualitative difference was observed—although total ice volume was approximately one-third less than in our supercontinent runs, and the bifurcation point for the transition between cold and ‘snowball’ conditions occurred at a higher CO_2 level, equivalent to about 1 W m^{-2} more infrared forcing. Runs with alternative Neoproterozoic reconstructions^{2,29} yielded similar results, although the critical CO_2 level again differed slightly ($<1\text{ W m}^{-2}$) for different geographies.

To determine what difference the ice sheet makes to our solution,

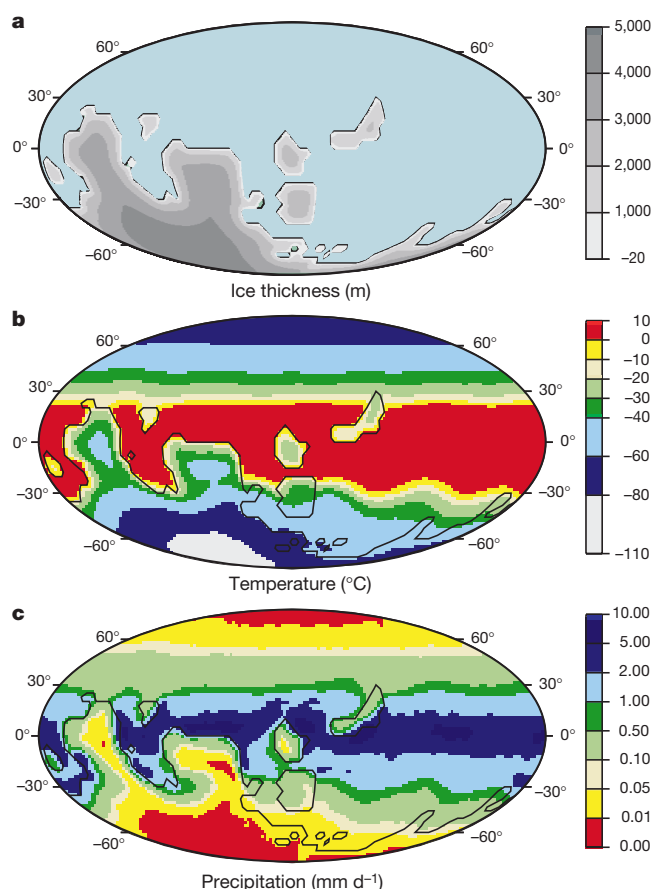


Figure 4 Late Precambrian (Neoproterozoic) climate simulations. **a**, Ice extent at equilibrium (Mollweide projection) for a ‘snowball Earth’ simulation. **b**, Annually averaged temperature for a GCM Neoproterozoic simulation using the above ice sheets as a lower boundary condition. A considerable area of open water exists in the tropics, though the ice sheets are almost everywhere below the melting point. **c**, Annually averaged precipitation in the GCM experiment. The cold temperatures and restricted area of open water result in a precipitation field which is less than 0.06 mm d^{-1} over the ice-sheet accumulation zone.

we ran the EBM with the same luminosity and CO_2 values used to obtain the ice-covered land-mass solution (note that snow and sea ice are still predicted). A comparison (Fig. 3) of the EBM alone and EBM/ice-model solutions indicates that an interactive ice sheet leads to a qualitatively different result. Ice elevation, and the long time constant associated with melting thick ice, allows the ice to ‘weather’ warm summer orbital configurations^{14,34} without deglaciating. Whereas ice growth is restricted by the time constants inherent in ice flow and precipitation, snow area can grow rapidly. A CO_2 increase equivalent to a $69\text{--}70\text{ W m}^{-2}$ warming is required to melt the ice (not shown). This is approximately equivalent to one-third of a bar of CO_2 . The deglaciation is very rapid, requiring less than 2,000 years to complete. Both of the above results are in agreement with ref. 4. This rapid and large warming would also explain the unusual configuration of tropical carbonates resting on diamictites—a characteristic feature of Neoproterozoic sediment sequences.

Although the coupled model generates a snowball Earth solution with sea-ice albedos greater than 0.5, the highly idealized sea-ice parameterization²⁵ raises questions about the validity of the oceanic component of the model simulation. In addition, the EBM’s diffusive heat transport mechanism is not as good a representation of tropical dynamics as it is of mid-latitude eddies. We therefore ran an additional set of stand-alone GCM mixed-layer ocean experiments (that is, without oceanic dynamics) with the fixed ice sheet obtained from our coupled run. We employed the GENESIS 2 GCM³⁵ which has diurnal forcing, interactive clouds, semi-lagrangian water vapour transport, and a six-layer sea-ice model. The sea-ice albedos vary between 0.7 and 0.8 (at visible wavelengths) and 0.4 to 0.5 (infrared) depending on whether temperatures are at the melting point, 5°C below the melting point, or in between. The total albedo at each ocean point is based on sea-ice fraction, itself dependent on the sea-ice thickness. The model was run at T31 resolution (approximately $3.75^{\circ} \times 3.75^{\circ}$). The initial conditions involved a zonally averaged state with an average temperature of 0°C . Simulations³⁶ used CO_2 values ranging from 0.5 to 2.5 times present levels, and the same luminosity decrease as in our coupled runs.

Simulations³⁶ with 0.5 and 1.0 times CO_2 (not shown) quickly yielded ice-covered oceans whereas a doubled CO_2 experiment resulted in ice-free tropical oceans but with a very slow drift to lower temperatures. We illustrate the last three years of the more stable $2.5 \times \text{CO}_2$ run (Fig. 4). An important feature of this experiment is that sea ice extends to only about 25° palaeolatitude (Fig. 4b), with sea-ice thickness varying from $\sim 1\text{ m}$ near the ice

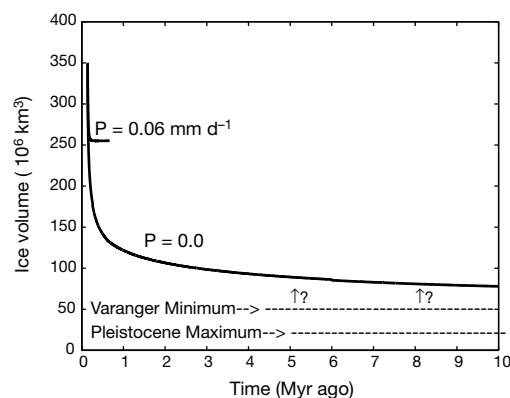


Figure 5 The effect of precipitation on global ice volume. When precipitation is reduced by a factor of ten, the ice sheets quickly achieve a slightly thinner equilibrium state. When precipitation is reduced to zero, the decline in ice volume is continuous. After 10 Myr the volume is about double that of the Pleistocene maximum, but comparable to some sea-level estimates⁴³ for the late Precambrian (Varanger) glaciation.

edge to 10 m in higher latitudes (the model cut-off limit). Open water can exist despite extensive ice sheets for two reasons—the $\sim 1,500$ -km length scale of the influence of ice sheets on temperatures^{37–39}, and an important negative feedback in the GCM that is not included in the EBM. Reduced absolute humidity in a colder world reduces cloudiness by about 35%, and hence the reduction in insolation due to the younger Sun is somewhat compensated by a lower equatorial albedo (compare ref. 40). Dynamical feedback between sea-ice formation and ocean circulation might further affect the equatorial extent of sea ice. Additional model simulations with the fully coupled climate/ice-sheet model (W. T. Hyde *et al.*, manuscript in preparation) mimicking this negative feedback yield open water in the tropics. However, in this open-water case isolated land masses in equatorial regions are ice-free. This problem will be examined further in subsequent work.

Even with this large area of open water zonally averaged precipitation over much of the ice sheet is less than 0.05 mm d^{-1} ; this is because most of the precipitation rains out over the open water, with very little transport into continental interiors. Temperatures below zero over most of the ice sheet indicate that open water is consistent with a mainly ice-covered land area (except for those ice sheets confined to small land masses in the equatorial ocean). Such open-water conditions near the ice are consistent with peritidal deposits⁴¹ for the south Australia (Fig. 1) glacial sequence.

The near-snowball climate, which was first conjectured by Kirschvink⁴², is qualitatively different from that of a true snowball Earth. This result is also not consistent with the interpretation⁴ of $\delta^{13}\text{C}$ records as indicating a near-abiotic ocean. However, if metazoans are found to have evolved by that time, and if the $\delta^{13}\text{C}$ record is shown to be open to alternative interpretation (see further comments below), then the open-water solution represents a straightforward reconciliation of these possibilities with equatorial glaciation.

Although the model can rather easily generate ice-covered continents with plausible changes in boundary conditions, the model-generated ice volume (nearly 10 times Pleistocene values) is in excess of estimated sea-level changes due to Precambrian glaciations^{4,7}. Such a change would also increase mean ocean salinities by about 20%. The very large ice sheet raises the question of whether the model produces reasonable ice volumes for a given ice area. Several lines of reasoning lead us to believe that our ice sheets are not unrealistically thick. For example, our ice sheets are considerably thinner than predicted by one published area/volume relation⁴⁴. The model's predicted Pleistocene ice volumes are in line with observations²¹. Although our single-domed Neoproterozoic simulation tends to exaggerate ice volume as compared with a more realistic multidomed reconstruction, a comparison between this model's single-domed Laurentide simulation²¹ and a multidomed simulation^{45,46} implies that this exaggeration is not large. The latter effect would also to some degree be offset by the fact that we do not allow ice to accumulate on continental shelves or epicontinental seas.

As our GCM results indicate (Fig. 4c), a radical reduction in precipitation should occur⁴ as a result of decreased temperatures and increased sea ice extent. However, the continued ablation of the ice sheet (largely due to calving of icebergs) would decrease ice volume. This would in turn allow sea level to rise, partially reversing the initial large sea-level drop. We tested this possibility through a series of runs in which the precipitation was lowered from 0.6 to 0.06 mm d^{-1} , and further to a limiting case of 0.0 mm d^{-1} , starting from the glaciated state of Fig. 4. The 10-Myr duration of our zero-precipitation run should be regarded as a minimum estimate for the time required to produce this sea-level drop, as the model uses a temperature- and pressure-independent rheology, and ice velocities would be much slower in the very cold snowball Earth, particularly as the ice sheets thinned, resulting in a slower rate of mass loss at the coast. When precipitation is reduced the ice sheets initially decay

quite rapidly, slowing as the ice surface slopes decrease (Fig. 5). A drop in precipitation by a factor of 10 decreases the equilibrium ice amount by about 30%, while a complete shut-off of moisture reduces the ice volume by 70%, but only after ~ 5 –10 million years. Although our 0.06 mm d^{-1} run predicts a sea-level decline far greater than a recent estimate⁴³ of 160 m for the Neoproterozoic, the magnitude of the observational constraint is still open to question. For example, it took more than a century to narrow down the sea-level decrease for the Last Glacial Maximum⁴⁷. The larger ice volume we simulate could explain the thickness of postglacial carbonates ($>300 \text{ m}$) precipitated on continental shelves. Such an amount could be accommodated by the degree of isostatic subsidence ($\sim 300 \text{ m}$) accompanying a 1,000-m ice sheet at the coast (Fig. 2a); overdeepening of continental shelves due to ice-sheet scouring would further enhance such accommodation (for reference, shelf depth is several hundred metres in places around Antarctica).

Discussion

An ice-covered Neoproterozoic land mass is predicted by our coupled climate/ice-sheet model with only two significant changes in boundary conditions from present values—a solar luminosity decrease consistent with estimates from solar physics, and a CO_2 level within about 50% of the present value. The final phase of the glacial advance, and the subsequent deglaciation, are consistent with the hypotheses proposed by Hoffman *et al.*⁴. These conclusions are relatively robust, in that they hold true for different land–sea configurations and continental freeboard. The results were obtained without any retuning of a model developed to explain the last glacial cycle.

A major finding of this study is that an area of open water in the equatorial oceans—which could have allowed for the survival of metazoans—is consistent with the evidence for equatorial glaciation at sea level. Although this result requires more model testing, it does not seem to be completely at variance with $\delta^{13}\text{C}$ data. Several studies suggest $\delta^{13}\text{C}$ values of -3‰ to -4‰ in deposits flanking the Varanger glaciation^{4,48,49}; such values are not consistent with a near-abiotic ocean, but would be consistent with the low level of ocean productivity that might have occurred in an ocean refugium of the size we simulate. Because of the need by early metazoans for a shallow water substrate, such an area could only be a true refugium if it contained ice-free continental shelves during the otherwise near-total glaciation. Identification of such sites would be a critical test of our open-water result (some of the equatorial land masses illustrated in Fig. 1 could have been in different locations 590 Myr ago). Length-scale arguments discussed above suggest that these sites would probably be isolated equatorial land masses at some distance from the main land mass (Fig. 1). If such sites are found, they should provide information about metazoa during one of the most significant 'bottlenecks' in the evolution of life. For example, selection pressure exerted on biota by the extreme climates may have led to rapid development of new forms during glaciation and afterwards, with the sea-level rise and expansion of biota into the large number of new and unpopulated habitats. Although there is clearly a need for more climate modelling and additional geological data (for example, clarification of the $\delta^{13}\text{C}$ and sea-level evidence, and more high-quality palaeomagnetic sites), our results indicate that inclusion of explicit ice-sheet physics significantly closes the gap between models and data for the largest glaciation of the past billion years and for one of the most critical intervals of evolution in Earth history. □

Received 9 December 1999; accepted 18 March 2000.

- Hoffman, P. F. Did the breakout of Laurentia turn Gondwanaland inside-out? *Science* **252**, 1409–1411 (1991).
- Dalziel, I. W. D. Overview: Neoproterozoic–Paleozoic geography and tectonics: Review, hypothesis, environmental speculation. *Geol. Soc. Am. Bull.* **109**, 16–42 (1997).
- Knoll, A. H. & Walter, M. R. Latest Proterozoic stratigraphy and Earth history. *Nature* **356**, 673–678 (1992).

4. Hoffman, P. F., Kaufman, A. J., Halverson, G. P. & Schrag, D. P. A Neoproterozoic snowball earth. *Science* **281**, 1342–1346 (1998).
5. Christie-Blick, N. Pre-Pleistocene glaciation on Earth: Implications for climatic history of Mars. *Icarus* **50**, 423–443 (1982).
6. Hambrey, H. A. & Harland, W. B. The late Proterozoic glacial era. *Paleogeogr. Paleoclimatol. Paleocol.* **51**, 255–272 (1985).
7. Eyles, N. Earth's glacial record and its tectonic setting. *Earth Sci. Rev.* **35**, 1–248 (1993).
8. Schmidt, P. W. & Williams, G. E. The Neoproterozoic climate paradox: Equatorial paleolatitude for Maninoan glaciation near sea level in south Australia. *Earth Planet. Sci. Lett.* **134**, 107–124 (1995).
9. Wray, G. A., Levontin, J. S. & Schapiro, L. H. Molecular evidence for deep Precambrian divergences among metazoan phyla. *Science* **274**, 568–573 (1996).
10. Breyer, J. A., Busbey, A. B., Hanson, R. E. & Roy, E. C. Possible new evidence for the origin of metazoans prior to 1 Ga: Sediment-filled tubes from the Mesoproterozoic Allamoore formation, Trans-Pecos Texas. *Geology* **23**, 269–272 (1995).
11. McNamara, K. J. Dating the origin of animals. *Science* **274**, 1995–1996 (1996).
12. Fedonkin, M. A. & Waggoner, B. M. The Late Precambrian fossil *Kimbrella* is a mollusc-like bilaterian organism. *Nature* **388**, 868–871 (1997).
13. North, G. R. Theory of energy-balance climate models. *J. Atmos. Sci.* **41**, 3990–3995 (1975).
14. Crowley, T. J. & Baum, S. K. Effect of decreased solar luminosity on late Precambrian ice extent. *J. Geophys. Res.* **98**, 16723–16732 (1993).
15. Jenkins, G. S. & Frakes, L. A. GCM sensitivity test using increased rotation rate, reduced solar forcing and orography to examine low latitude glaciation in the Neoproterozoic. *Geophys. Res. Lett.* **25**, 3528 (1998).
16. Jenkins, G. S. & Smith, S. R. GCM simulations of Snowball Earth conditions during the late Proterozoic. *Geophys. Res. Lett.* **26**, 2263–2266 (1999).
17. Williams, G. E. Late Precambrian glacial climate and the Earth's obliquity. *Geol. Mag.* **112**, 441–465 (1975).
18. Ogleby, R. J. & Ogg, J. G. The effect of large fluctuations in the obliquity on climates of the late proterozoic. *Paleoclim. Data Modelling* **2**, 293–316 (1998).
19. Deblonde, G. & Peltier, W. R. Simulations of continental ice sheet growth over the last glacial-interglacial cycle: Experiments with a one-level seasonal energy balance model including realistic geography. *J. Geophys. Res.* **96**, 9189–9215 (1991).
20. Deblonde, G., Peltier, W. & Hyde, W. T. Simulations of continental ice-sheet growth over the glacial-interglacial cycle: Experiments with a one level seasonal energy balance model including seasonal ice albedo feedback. *Glob. Planet. Change* **98**, 37–55 (1992).
21. Tarasov, L. & Peltier, W. R. Terminating the 100 kyr ice age cycle. *J. Geophys. Res. D* **18**, 21665–21693 (1997).
22. Nye, J. F. The motion of ice sheets and glaciers. *J. Glaciol.* **3**, 493–507 (1959).
23. Reeh, N. Parameterization of melt rate and surface temperature on the Greenland ice sheet. *Polarforschung* **59**, 113–128 (1990).
24. Huybrechts, P. & T'Siobbel, S. Thermomechanical modelling of Northern Hemisphere ice sheets with a two-level mass-balance parameterization. *Ann. Glaciol.* **21**, 111–116 (1995).
25. Hyde, W. T., Kim, K.-Y., Crowley, T. J. & North, G. R. On the relationship between polar continentality and climate: Studies with a nonlinear energy balance model. *J. Geophys. Res. D* **11**, 18653–18668 (1990).
26. Crowley, T. J., Baum, S. K. & Hyde, W. T. Climate model comparisons of Gondwanan and Laurentide glaciations. *J. Geophys. Res. D* **5**, 619–629 (1999).
27. Hyde, W., Crowley, T. J., Tarasov, L. & Peltier, W. R. The Pangean ice age: Studies with a coupled climate-ice sheet model. *Clim. Dyn.* **15**, 619–629 (1999).
28. Berger, A. Long-term variations of daily insolation and Quaternary climate changes. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
29. Bond, G. C., Nickeson, P. A. & Kominz, M. A. Breakup of a supercontinent between 623 and 555 Ma: New evidence and implications for continental histories. *Earth Planet. Sci. Lett.* **70**, 325–345 (1984).
30. Kaufman, A. J., Jacobson, S. B. & Knoll, A. H. The Vendian record of Sr and C isotopic variations in seawater: Implications for tectonics and paleoclimate. *Earth Planet. Sci. Lett.* **120**, 409–430 (1993).
31. Baum, S. K. & Crowley, T. J. Seasonal snowline instability in a climate model with realistic geography: Application to Carboniferous (≈ 300 Ma) glaciation. *Geophys. Res. Lett.* **18**, 1719–1722 (1991).
32. Crowley, T. J., Yip, K.-J. J. & Baum, S. K. Snowline instability in a general circulation model: Application to Carboniferous glaciation. *Clim. Dyn.* **10**, 363–376 (1994).
33. Scotese, C. R. & Golonka, J. *Paleogeographic Atlas* (Technical report, Paleomap project, University of Texas-Arlington, 1992).
34. Crowley, T. J., Yip, K.-J. J. & Baum, S. K. Milankovitch cycles and Carboniferous climate. *Geophys. Res. Lett.* **20**, 1175–1178 (1993).
35. Thompson, S. L. & Pollard, D. A global climate model (GENESIS) with a land-surface-transfer scheme (LSX). Part 1: Present-day climate. *J. Clim.* **8**, 732–761 (1995).
36. Baum, S. K. & Crowley, T. J. GCM response to Late Precambrian (~ 600 Ma) ice-covered continents. *J. Geophys. Res.* (submitted).
37. North, G. R. The small ice cap instability in diffusive climate models. *J. Atmos. Sci.* **32**, 1301–1307 (1984).
38. Manabe, S. & Broccoli, A. J. The influence of continental ice sheets on the climate of an ice age. *J. Geophys. Res.* **90**, 2167–2190 (1985).
39. Hyde, W. T., Crowley, T. J., Kim, K.-Y. & North, G. R. Comparison of GCM and energy balance model simulations of seasonal temperature changes over the past 18000 years. *J. Clim.* **2**, 864–887 (1989).
40. Jenkins, G. S. A general circulation model study of the effects of faster rotation rate, enhanced CO₂ concentration and reduced solar forcing: Implications for the faint-young sun paradox. *J. Geophys. Res.* **98**, 20803–20811 (1993).
41. Williams, G. E. Cyclicity in the Late Precambrian Elatina formation, South Australia: Solar or tidal signature? *Clim. Change* **13**, 117–128 (1998).
42. Kirschvink, J. L. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 51–52 (Cambridge University Press, 1992).
43. Christie-Blick, N. in *Proterozoic to Recent Stratigraphy, Tectonics and Volcanology, Utah, Nevada, Southern Idaho and Central Mexico* (eds Link, P. K. & Kowallis, B. J.) 1–30 (Geology Studies Vol. 42, Part I, Brigham Young University, Provo, 1997).
44. Paterson, W. S. B. *The Physics of Glaciers* (Pergamon, Oxford, 1981).
45. Peltier, W. R. Postglacial variations in the level of the sea: Implications for climate dynamics and solid-earth geophysics. *Rev. Geophys.* **36**, 603–689 (1998).
46. Tarasov, L. & Peltier, W. R. Impact of thermo-mechanical ice-sheet coupling on a model of the 100 kyr ice age cycle. *J. Geophys. Res.* **104**, 9517–9545 (1999).
47. Fairbanks, R. G. A 17,000-year glacio-eustatic sea level record: Influence of glacial melting rates on Younger Dryas event and deep-ocean circulation. *Nature* **342**, 637–642 (1989).
48. Kaufmann, A. J., Knoll, G. H. & Narbonne, G. M. Isotopes, ice ages, and terminal Proterozoic Earth history. *Proc. Natl Acad. Sci. USA* **94**, 6600–6605 (1997).
49. Kennedy, M. J., Runnegar, B., Prave, A. R., Hoffman, K. H. & Arthur, M. A. Two or four Neoproterozoic glaciations? *Geology* **26**, 1059–1063 (1998).

Acknowledgements

We thank I. Dalziel and L. Gahagan for the plate reconstructions, and N. Christie-Blick, N. Eyles, I. Dalziel, P. Hoffman, M. Huber, A. Knoll and D. Schrag for comments and suggestions. We thank P. Smith for helping to prepare the cover picture. This research was supported by the National Center for Atmospheric Research, the National Sciences and Engineering Research Council of Canada, and by the NSF.

Correspondence and requests for materials should be addressed to W.T.H. (e-mail: hyde@rossby.tamu.edu).

A triplet of differently shaped spin-zero states in the atomic nucleus ^{186}Pb

A. N. Andreyev*, M. Huyse*, P. Van Duppen*, L. Weissman*, D. Ackermann†, J. Gerl†, F. P. Heßberger†, S. Hofmann†, A. Kleinböhl†, G. Münzenberg†, S. Reshitko†, C. Schlegel†, H. Schaffner†, P. Cagarda‡, M. Matos‡, S. Saro‡, A. Keenan§, C. Moore§, C. D. O'Leary§, R. D. Page§, M. Taylor§, H. Kettunen||, M. Leinoll, A. Lavrentiev¶, R. Wyss# & K. Heyde*

* Instituut voor Kern- en Stralingsfysica, University of Leuven, Celestijnenlaan 200 D, B-3001 Leuven, Belgium

† Gesellschaft für Schwerionenforschung Darmstadt, Postfach 110541, D-6100 Darmstadt, Germany

‡ Department of Nuclear Physics, Comenius University, Bratislava, Slovakia

§ Department of Physics, Oliver Lodge Laboratory, University of Liverpool, Liverpool L69 7ZE, UK

|| Department of Physics, University of Jyväskylä, FIN-40351 Jyväskylä, Finland

¶ Flerov Laboratory of Nuclear Reactions, Joint Institute for Nuclear Research, 141980 Dubna, Russia

Department of Physics, Royal Institute of Technology, 104 05 Stockholm, Sweden, and Department of Technology, Kalmar University, Box 905, 391 29 Kalmar, Sweden

* Vakgroep Subatomaire en Stralingsfysica, Institute for Theoretical Physics, B-9000 Gent, Belgium

Understanding the fundamental excitations of many-fermion systems is of significant current interest. In atomic nuclei with even numbers of neutrons and protons, the low-lying excitation spectrum is generally formed by nucleon pair breaking and nuclear vibrations or rotations. However, for certain numbers of protons and neutrons, a subtle rearrangement of only a few nucleons among the orbitals at the Fermi surface can result in a different elementary mode: a macroscopic shape change^{1–3}. The first experimental evidence for this phenomenon came from the observation of shape coexistence in ^{16}O (ref. 4). Other unexpected examples came with the discovery of fission isomers⁵ and super-deformed nuclei⁶. Here we find experimentally that the lowest three states in the energy spectrum of the neutron deficient nucleus ^{186}Pb are spherical, oblate and prolate. The states are populated by the α -decay of a parent nucleus; to identify them, we combine knowledge of the particular features of this decay⁷ with sensitive measurement techniques (a highly efficient velocity filter⁸ with strong background reduction, and an extremely selective recoil- α -electron coincidence tagging method^{8–10}). The existence of this apparently unique shape triplet is permitted only by the specific conditions that are met around this particular nucleus.

The study of a system of A nucleons bound together by nucleon–nucleon interaction in an atomic nucleus poses a complex problem. The rich spectrum of elementary modes of collective motion, such as rotations and vibrations, is very difficult to derive from the action of the nucleon–nucleon force. Instead, these phenomena can be described as emerging from the interplay of the constituent particles and the use of effective forces. In this context, the atomic nucleus which, at the lower energy scale is now understood as behaving very much like a superfluid system, can be compared to complex mesoscopic systems of the type that appears in atomic cluster structures and low-dimensional quantum systems (integer and fractional quantum Hall effect). This is due to the number of nucleons (of the order of 100 in an average nucleus), which is too large for a full microscopic study but also too small for statistical treatments. The variety of shapes in atomic nuclei and the effect of individual nucleons continue to be a topic of current research (see ref. 11 for a review).

The lowest excitation mode encountered in atomic nuclei is, in

general, associated with rotations. In the mass $A = 170$ region, the observed moment of inertia (J) of a deformed nucleus lies between the values corresponding to a rigid object and a fluid, and leads to a rotational energy constant of the order of $\hbar^2/2J \approx 15$ keV. But because some nucleon numbers are found to be ‘magic’ (N or $Z = 8, 20, 28, 50, 82, 126$ (or $Z \approx 114$)), nuclei with a closed shell are predominantly spherical and as such do not exhibit rotational excitations at low energy. Due to the stiffness of the closed configuration, vibrations will also be hindered and the first excitations will be due to breaking a pair of nucleons from the open shell. The resulting spectrum is a typical one-broken-pair multiplet and leads in the even–even Pb ($Z = 82$) nuclei to a first excited state around 1 MeV with spin and parity 2^+ , except for doubly magic ^{208}Pb which has a 3^- level at 2.6 MeV as its first excited state.

A new phenomenon occurs in the neutron-deficient Pb nuclei, where the number of valence neutrons becomes maximal as the neutron number lies in the middle of two numbers characteristic of closed shells ($N = 82$ and $N = 126$)^{1,2}. Within a shell-model approach², a proton pair excitation across the $Z = 82$ shell leads to the creation of two valence proton particles (2p) above, and two valence proton holes (2h) below, the $Z = 82$ closed shell, that are coupled to spin and parity 0^+ . The energy cost is about twice the energy of the $Z = 82$ shell gap ($\approx 2 \times 3.9$ MeV = 7.8 MeV) but the creation of one particle and one hole pair yields a gain in pairing correlation energy of 2.6 MeV which can be derived from the experimental one- and two-proton separation energies². Furthermore, the proton shell is now broken up and the attractive proton–neutron interaction will lower the energy for these 2p–2h states, so that from ^{194}Pb onwards this 0^+ state lies below the first excited 2^+ state³. Within a deformed mean-field approach this is equivalent to a macroscopic phase change from a spherical to an oblate shape¹². The band structure built on top of it has been observed only for a limited number of neutron-deficient Pb nuclei, the lightest being

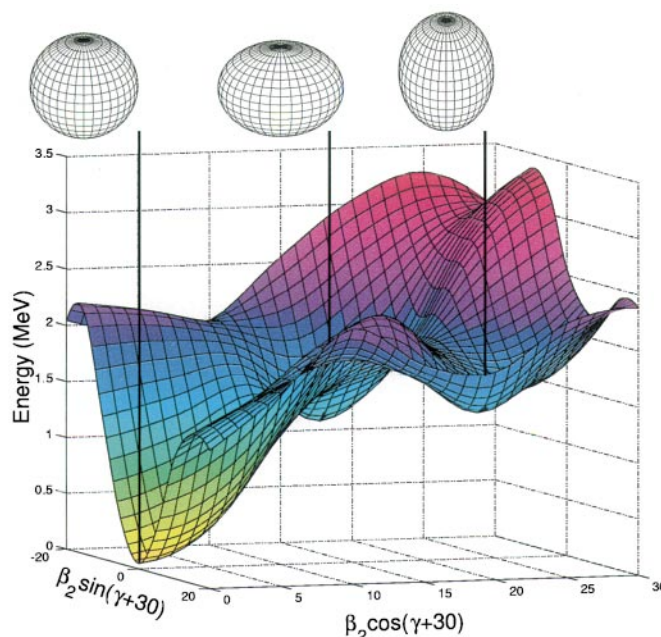


Figure 1 Calculated potential energy surface of ^{186}Pb . Spherical, oblate and prolate minima are indicated by thick vertical black lines. Calculations are performed on the cartesian mesh. The β_2 parameter expresses the elongation of the nucleus along the symmetry axis, while the γ parameter relates to the degree of triaxiality in the deformation (that is, the relative lengths of the three principal axes of the spheroidal nucleus). The γ parameter is defined such that $\gamma = 0$ corresponds to a prolate (cigar-like) shape and $\gamma = 60$ to an oblate (disk-like) shape.

^{190}Pb (ref. 13). This schematic picture can also be extended to multi-particle and multi-hole excitations, each of them possibly leading to a different shape.

The energy of the different shape configurations can be calculated using a nucleon potential, such as the Woods–Saxon¹⁴ potential, in which the energy of a single-particle orbital depends on the deformation. These potential-energy-surface calculations have become more and more sophisticated with time, resulting in accurate descriptions of the nuclear shapes and the configurations involved^{14–16}. Figure 1 gives the result of such a calculation for ^{186}Pb and shows, next to the spherical minimum corresponding to the ground state, an oblate and a prolate minimum at an excitation energy of around 1 MeV. In the particle–hole picture, the former corresponds to a 2p–2h configuration, while the latter corresponds to a 4p–4h configuration as discussed in ref. 15. It is of prime importance to identify and to characterize the three 0^+ states associated with the three different shapes.

Alpha decay forms an ideal tool to populate the 0^+ bandheads and probe selectively the underlying structure⁷; this has led to the identification of the oblate 2p–2h bandheads down to $^{188}\text{Pb}_{106}$, where a bandhead was observed at $E^* = 591$ keV (ref. 9, 17, 18) above the spherical ground state and was recently confirmed by in-beam electron spectroscopy¹⁹. In-beam studies showed rotational bands in $^{184,186,188}\text{Pb}$ associated only with a prolate-deformed shape^{20–22}. Conflicting results on the prolate 4p–4h 0^+ bandhead in ^{188}Pb have been reported, giving two different values of 767(12) keV (ref. 9) (α -decay study) and 725(4) keV (ref. 19) (conversion-electron spectroscopy). (Here 767(12) keV, for example, indicates 767 ± 12 keV.) The exotic nucleus ^{186}Pb , far away from the stability line but with a closed proton shell and with the neutrons exactly at mid-shell between $N = 82$ and $N = 126$, represents a unique opportunity to study the competition between shape excitations, collective motion and intrinsic excitations. This maximum in proton–neutron correlations will enable the study of shape coexistence at the lowest possible excitation energy. The relative positions of the three different shapes, the height of the barrier between them, and consequently the purity of the configurations, are further aspects that can be addressed in this nucleus.

To study in detail ^{186}Pb through the α -decay of ^{190}Po is a challenging task. The ^{190}Po nuclei are produced via the 4n-evaporation channel for the fusion reaction of 200-pnA ($= 1.2 \times 10^{12}$ particles per s) 255-MeV ^{52}Cr ions with a

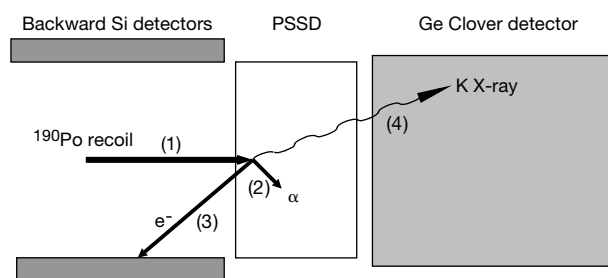


Figure 2 A view of the detector system (not to scale). Owing to the segmentation of the position sensitive silicon detector (PSSD) and the low implantation rate, it is possible to correlate the time and position of the implanted recoil products and their subsequent α -decays (recoil- α correlations). Six Si detectors (the Si box) are mounted in front of the PSSD ('backward') to detect conversion electrons in prompt coincidence with the α -decay (recoil- α -e $^-$ correlations). A 4-fold segmented Ge Clover detector records prompt α -X and α - γ coincidences. A schematic picture of a recoil (1) – α (2) – e $^-$ (3) – X (4) event is given. The identification of the different α lines is based on the half-life, the count rate as a function of the bombarding energy (5 different beam energies), coincidences with electrons, characteristic K-X rays and γ lines, correlations with known α -decay of daughters in the decay chain, and sum-energy relations.

$290 \mu\text{g cm}^{-2}$ $^{142}\text{NdF}_3$ target. The cross-section is only about 300 nbarn ($= 3 \times 10^{-31} \text{ cm}^2$) which corresponds to about 300 ^{190}Po atoms being produced per hour but with a background of fission and transfer reactions about a million times higher. The Separator for Heavy Ion Production (SHIP) velocity filter⁸, developed for the studies of superheavy elements, is able to separate the nuclei of interest from unwanted background and to implant them into an efficient detection system^{8,10} (see Fig. 2 and its legend for further details). The experimental set-up has been optimized to observe fine structure in the α -decay that leads, as studied for the heavier even–even Pb nuclei⁷, to the identification of low-lying 0^+ bandheads, which will decay predominantly by E0 conversion electron transitions to the ground state.

The data in Fig. 3 form the main evidence for the observation of three different shapes in ^{186}Pb . Out of the 27,080 counts recorded in Fig. 3a, only 274 show (see Fig. 3b) coincidence with electrons, as observed from the two peak structures at 7,012(20) and 6,896(20) keV. As seen in Fig. 3c, these peaks are almost completely decayed within 12 ms. The half-life of the 7,012-keV line is 2.6(0.3) ms, while for the 6,896-keV line an upper limit of 5 ms for the short-lived component is obtained. Only two isotopes shown in Fig. 3a have such a short half-life: ^{190}Po , 2.45(5) ms (this work); and $^{189\text{m}}\text{Bi}$, 5.0(1) ms (ref. 10). All the other isotopes have considerably longer half-lives (see Fig. 3a). The behaviour of the $^{189\text{m}}\text{Bi}$ isotope was studied at different beam energies, and it can be concluded that α -e coincidences from fine structure in the α -decay

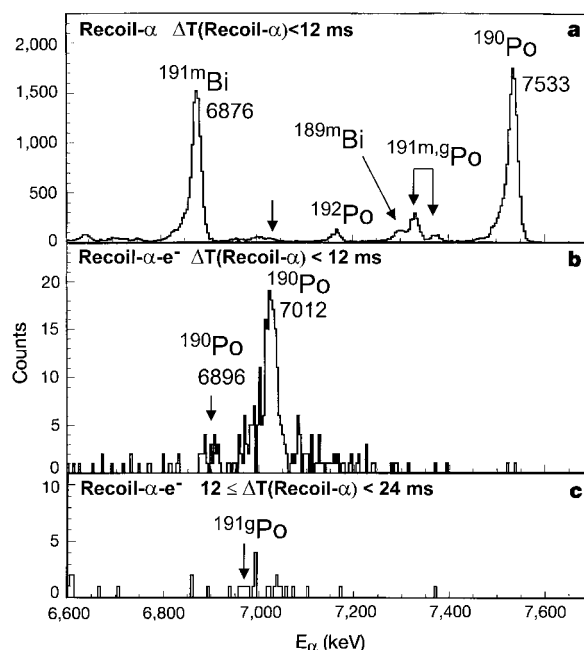


Figure 3 Three α spectra with different gating conditions. **a**, Part of the α spectrum recorded in the PSSD within 12 ms after the recoil implantation. Some peaks are labelled with the α -decay energy and isotope they relate to, with the label g for the ground state and m for an isomeric state for the same isotope. The strongest α line is the ground-state to ground-state α transition of ^{190}Po at 7,533(10) keV decaying with a half-life of 2.45(5) ms, an improved value compared to literature^{18,23}. The other α lines are from isotopes produced in other open evaporation channels (xn, pxn). They have been identified on the basis of their energy, half-life and yield as a function of the beam energy (excitation function). The main isotopes are: $^{191\text{m}}\text{Bi}$, $T_{1/2} = 115(10)$ ms (ref. 24); ^{192}Po , $T_{1/2} = 33.4(14)$ ms (ref. 17); $^{189\text{m}}\text{Bi}$, $T_{1/2} = 5.0(1)$ ms (ref. 10); $^{191\text{m}}\text{Po}$, $T_{1/2} = 98(8)$ ms (ref. 25); $^{191\text{g}}\text{Po}$, $T_{1/2} = 22(1)$ ms (ref. 25). **b**, The same as **a**, but in a prompt coincidence with conversion electrons, registered in the backward detectors. The random background probability is 7×10^{-5} . **c**, The same prompt electron condition as in **b**, but now with α events that are registered between 12 and 24 ms after implantation.

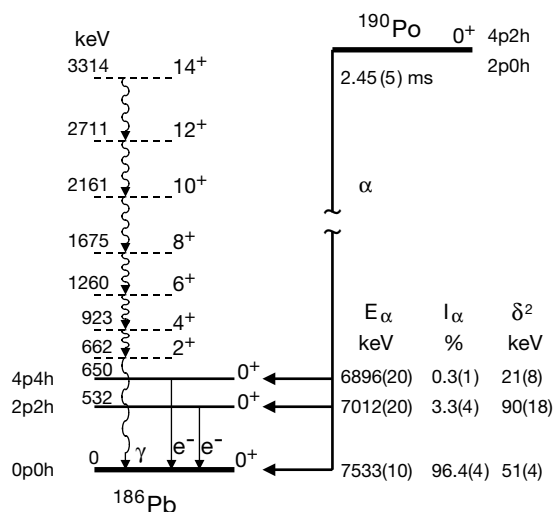


Figure 4 The decay pattern of ^{190}Po and the level scheme of ^{186}Pb . Indicated are α -decay energies E_α , intensities I_α , reduced α -widths δ^2 , and configuration assignments. As discussed in ref. 7, the reduced α widths δ^2 for the even-mass Po, Pb, Hg and Pt nuclei around $N = 104$ lie around 70 keV. Lower (higher) values indicate retarded (enhanced) transitions. The known prolate rotational band ($I^\pi = 2^+ \rightarrow 14^+$) is shown by dashed lines^{21,22}.

of $^{189\text{m}}\text{Bi}$ (ref. 10) do not contribute to the spectrum in Fig. 3b. The only coincidences that the two lines observed in Fig. 3b show with Ge detectors are in the region of low-energy X-rays below 100 keV, and the 7,012-keV α line is clearly coincident with the characteristic K X-rays of Pb. Furthermore, both lines correlate with the α line of ^{186}Pb , the daughter of ^{190}Po . Therefore the 7,012 and 6,896 keV α lines are attributed to the decay of ^{190}Po . The counts in the region between 6,800 and 7,000 keV in Fig. 3c) can be explained by coincidences with conversion electrons from electromagnetic transitions, recently reported in the decay of $^{191\text{m,g}}\text{Po}$ (refs 24, 25). Their decay is partly responsible for the left shoulder of the 7,012-keV peak observed in Fig. 3b, and for the weak long-lived component under the 6,896-keV line. The extra counts above the 7,012-keV peak (Fig. 3b) are due to an electron sharing its energy between the position-sensitive silicon detector and the Si box, thus partially adding to the α line. By comparing the coincidences of the 7,012 and 6,896 keV α lines with electrons, X- and γ -rays, the only valid conclusion is to assign the two lines as fine structure lines due to α -decay of ^{190}Po feeding excited levels in ^{186}Pb , subsequently decaying by E0 conversion electron emission to the ground state (Fig. 4).

From the reduced α -decay widths δ^2 of $\Delta I = 0$ transitions, which correspond to the transition rate after correcting for the energy dependence from the tunnelling through the Coulomb barrier, it is possible to obtain information on the internal structure of the configurations involved in the connected states^{7,26,27}. This method has been used to show that the ground state of ^{186}Pb is spherical, as are all heavier even-even Pb nuclei^{26,28}. The ground states of even-even Po nuclei with $A > 196$ exhibit a spherical 2p proton configuration. However, owing to the rapid lowering in energy of deformed 4p–2h proton configurations, the ground states of the lighter even-even Po nuclei, from ^{196}Po onwards²⁷, develop a mixed character of 2p and 4p–2h configuration (with $\sim 70\%$ of the latter in the ground state of ^{192}Po ; refs 9, 17). The reduced widths in the α -decay of ^{190}Po (see Fig. 4) and ^{192}Po corroborate this superposition, and unambiguously show that the decay to the 0p–0h spherical ground state in ^{186}Pb is much slower than the decay to the first excited 0^+ state. This indicates that the ground state of ^{190}Po is mainly the deformed 4p–2h configuration, and that the first excited 0^+ state in ^{186}Pb has a multiparticle–multihole character. Based on the energy systematic of the oblate bandheads in the heavier Pb

isotopes, the extrapolation of the 0^+ bandhead energy starting from the higher-spin members of the prolate band in ^{186}Pb (see Fig. 4), and the reduced α -widths, we identify the 0^+ state (with an excitation energy of 532 keV) as the oblate state, and the 0^+ state (with an excitation energy of 650 keV) as the prolate state. Both of these levels are below the first excited 2^+ state, at 662 keV.

This observation allows to consider the atomic nucleus ^{186}Pb as a ‘laboratory’ in which to observe the effect of a change of a small number of particles on the macroscopic shape. The energy gap at the $Z = 82$ closed shell (~ 3.9 MeV) forces the protons to adopt a spherical shape, while the large space available for the valence neutrons (with N values between 82 and 126) favours a deformed configuration. Combined with the strength of the proton–neutron quadrupole force and the pairing², a subtle balance originates where the promotion of specific proton pairs to the next shell leads to either a prolate or to an oblate state. Proceeding to lighter masses (towards Sn and Ni), the steady increase in the shell gap energy, combined with a decreasing open space for the valence particles, lead to an increase in excitation energy of the deformed states. For the heavier elements (beyond Pb), the shell gap becomes too small and the proton–neutron correlation energy causes the ground state to become deformed. Only in the neutron-deficient Pb region are the exact conditions met to study, near the ground state (in a zone of low-level density), the microscopic origin of the spherical-symmetry breaking which leads to spheroidal deformation. Other systems having the character of Fermi liquids, such as atomic and molecular clusters, also exhibit shell structure and deformation. However, the large moments of inertia make rotation the easiest way by far to excite the system ($\hbar^2/2J \approx 10^{-4}$ eV) thereby complicating the observation of shape changes.

The study presented here has used α -decay as a spectroscopic tool to elucidate this triple shape coexistence. Nuclear physics has developed a whole arsenal of spectroscopic methods to identify and characterize the collective behaviour of quantum liquids. The observation of band structures build upon different shapes and their interaction will further probe the purity of the states and their evolution when excited. □

Received 20 December 1999; accepted 3 April 2000.

- Heyde, K. *et al.* Coexistence in odd-mass nuclei. *Phys. Rep.* **102**, 291–393 (1983).
- Wood, J. L. *et al.* Coexistence in even-mass nuclei. *Phys. Rep.* **215**, 101–201 (1992).
- Van Duppen, P. *et al.* Observation of low-lying $I^\pi = 0^+$ states in the single-closed shell nuclei $^{192-198}\text{Pb}$. *Phys. Rev. Lett.* **52**, 1974–1977 (1984).
- Morinaga, H. Interpretation of some excited states of 4n self-conjugate nuclei. *Phys. Rev.* **101**, 254–258 (1956).
- Björnholm, S. & Lynn, J. E. The double-humped fission barrier. *Rev. Mod. Phys.* **52**, 725–931 (1980).
- Nolan, P. & Twin, P. Superdeformed states at high angular momentum. *Annu. Rev. Nucl. Part. Sci.* **38**, 533–562 (1988).
- Van Duppen, P. & Huyse, M. Shape coexistence around the $Z = 82$ closed shell probed by α decay. *Hyperfine Interact.* (in the press).
- Hofmann, S. New elements—approaching $Z = 114$. *Rep. Prog. Phys.* **69**, 827–888 (1998).
- Allatt, R. G. *et al.* Fine structure in ^{192}Po α -decay and shape coexistence in ^{186}Pb . *Phys. Lett. B* **437**, 29–34 (1998).
- Andreyev, A. N. *et al.* Alpha decay of the new isotopes $^{188,189}\text{Po}$. *Eur. Phys. J. A* **6**, 381–385 (1999).
- Walker, P. & Dracoulis, G. D. Energy traps in atomic nuclei. *Nature* **399**, 35–40 (1999).
- Heyde, K. *et al.* Equivalence of the spherical and deformed shell-model approach to intruder states. *Phys. Lett. B* **218**, 287–290 (1989).
- Dracoulis, G. D. *et al.* Interference between spherical and deformed states in ^{190}Pb . *Phys. Lett. B* **432**, 37–44 (1998).
- Cwiok, S. *et al.* Axially deformed Woods-Saxon potential. *Comput. Phys. Commun.* **46**, 379–399 (1987).
- Nazarewicz, W. Variety of shapes in the mercury and lead isotopes. *Phys. Lett. B* **305**, 195–201 (1993).
- Tajima, N. *et al.* Diabatic effects in ^{186}Pb : a generator-coordinate analysis. *Nucl. Phys. A* **551**, 409–433 (1993).
- Bijnens, N. *et al.* Fine structure in the α decay of ^{192}Po . *Z. Phys. A* **356**, 3–4 (1996).
- Andreyev, A. N. *et al.* α -decay characteristics of neutron-deficient $^{190,192}\text{Po}$ nuclei and alpha branching ratios of $^{186,188}\text{Pb}$. *J. Phys. G* **25**, 835–837 (1999).
- Le Coz, Y. *et al.* Evidence of multiple shape coexistence in ^{186}Pb . *Eur. Phys. J. Direct A* **3**, 1–6 (1999).
- Cocks, J. F. C. *et al.* First observation of excited states in ^{184}Pb : spectroscopy beyond the neutron mid-shell. *Eur. Phys. J. A* **3**, 17–20 (1998).
- Heese, J. *et al.* Evidence for low-lying prolate bands in ^{188}Pb and ^{186}Pb . *Phys. Lett. B* **302**, 390–395 (1993).
- Baxter, A. M. *et al.* Spectroscopy of ^{186}Pb with mass identification. *Phys. Rev. C* **48**, R2140–R2143 (1993).

23. Batchelder, J. *et al.* α -decay properties of ^{190}Po and identification of ^{191}Po . *Phys. Rev. C* **55**, R2142–R2145 (1997).
24. Andreyev, A. *et al.* Fine structure in the α decay of the neutron-deficient ^{191}Po and ^{191}Bi isotopes. *Acta Phys. Pol.* **30**, 1255–1262 (1999).
25. Andreyev, A. *et al.* Hindered ($\Delta I = 0$) α decay and shape staggering in ^{191}Po . *Phys. Rev. Lett.* **82**, 1819–1822 (1999).
26. Wauters, J. *et al.* Fine structure in the α decay of even-even nuclei as an experimental proof for the stability of the $Z = 82$ magic shell at the very neutron-deficient side. *Phys. Rev. Lett.* **72**, 1329–1332 (1994).
27. Bijmens, N. *et al.* Intruder states and the onset of deformation in the neutron-deficient even-even polonium isotopes. *Phys. Rev. Lett.* **75**, 4571–4574 (1995).
28. Wauters, J. *et al.* α decay of ^{186}Pb and ^{184}Hg : the influence of mixing of 0^+ states on α -decay transition probabilities. *Phys. Rev. C* **50**, 2768–2773 (1994).

Acknowledgements

A.N.A. is on leave from FLNPh, JINR, Dubna, Russia. We are indebted to the UNILAC staff for the excellent performance of the ion source and the accelerator. We thank the colleagues from the GSI target laboratory for the preparation of the large area $^{142}\text{NdF}_3$ targets. This work was supported by the Access to Large Scale Facility programme under the Training and Mobility of Researchers programme of the European Union, by the FWO–Vlaanderen and the GOA programme, by UK EPSRC, by the Academy of Finland and by the Swedish Natural Research Council (NFR).

Correspondence and requests for materials should be addressed to M.H. (e-mail: mark.huyse@fys.kuleuven.ac.be).

Microdomain patterns from directional eutectic solidification and epitaxy

Claudio De Rosa^{*†}, Cheolmin Park^{‡†}, Edwin L. Thomas[‡] & Bernard Lotz[§]

^{*} Dipartimento di Chimica, Università di Napoli, “Federico II” Via Mezzocannone 4, 80134 Napoli, Italy

[‡] Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[§] Centre de Recherches sur les Macromolécules, CNRS Strasbourg, 67083 France

[†] These authors contributed equally to this work

Creating a regular surface pattern on the nanometre scale is important for many technological applications, such as the periodic arrays constructed by optical microlithography that are used as separation media in electrophoresis¹, and island structures used for high-density magnetic recording devices². Block copolymer patterns can also be used for lithography on length scales below 30 nanometres (refs 3–5). But for such polymers to prove useful for thin-film technologies, chemically patterned surfaces need to be made substantially defect-free over large areas, and with tailored domain orientation and periodicity. So far, control over domain orientation has been achieved by several routes^{6–9}, using electric fields, temperature gradients, patterned substrates and neutral confining surfaces. Here we describe an extremely fast process that leads the formation of two-dimensional periodic thin films having large area and uniform thickness, and which possess vertically aligned cylindrical domains each containing precisely one crystalline lamella. The process involves rapid solidification of a semicrystalline block copolymer from a crystallizable solvent between glass substrates using directional solidification and epitaxy. The film is both chemically and structurally periodic, thereby providing new opportunities for more selective and versatile nanopatterned surfaces.

Block copolymers consist of chemically distinct macromolecules covalently linked to form a single chain; due to their mutual repulsion, the dissimilar blocks tend to segregate into different domains whose shape, size and spacing are determined by the

relative amount of the block components and their respective molecular masses^{10,11}. Control over the microdomain patterns has been achieved by employing electric fields, temperature gradients, patterned substrates and neutral confining surfaces. For example, a temperature gradient has been used¹² to produce a vertically aligned lamellar structure with excellent long-range order. To obtain a well ordered structure, slow growth (millimetres per day) in a large gradient (700 °C per centimetre) is required.

In crystalline materials, control of the solidification process is central to many technologies that rely on the features of the resultant microstructure for achieving optimum properties. For example, the directional solidification of a eutectic metal alloy can lead to rod or lamellar structures well aligned along the growth direction¹³. In crystalline polymeric materials, orientation of crystallizable macromolecules has been achieved by mechanical forces, as in fibre spinning and also by epitaxial crystallization onto substrates^{14,15}. A vertically oriented lamellar structure has been achieved¹⁶ by utilizing crystallization from the microphase-separated state of a low-molecular-mass diblock copolymer adjacent to boundaries formed by dewetting from the substrate. Mixtures of polyethylene and a crystallizable solvent (tetrachlorobenzene) form a binary eutectic¹⁷, and polyethylene homopolymer has been shown¹⁸ to epitaxially crystallize on benzoic acid. We also note that benzoic acid and paraffin have been shown¹⁹ to form a eutectic, and that the paraffin grows epitaxially onto the crystallized benzoic acid. We thus considered that a semicrystalline block copolymer dissolved in a crystallizable and epitaxy-forming solvent might yield interesting directionally solidified structures.

We used a polystyrene-block-polyethylene (PS-PE) diblock copolymer that was prepared by hydrogenation of polystyrene-block-1,4-butadiene, previously synthesized via sequential anionic polymerization. The amorphous PS block and the crystallizable PE block have molecular masses of 40,000 and 10,000 daltons respectively, the volume fraction of the PE block is 0.24, and the melting point of the PE block is 98 °C. Small-angle X-ray scattering measurements of bulk films of PS-PE shows multiple low-angle reflections characteristic of hexagonally packed PE cylinders with the first Bragg peak at $q = 0.15 \text{ nm}^{-1}$, corresponding to a cylinder–cylinder spacing of 42 nm. (Here $q = (4\pi/\lambda)\sin\theta$, where λ is the X-ray wavelength and θ is $1/2$ of the scattering angle.)

When cast from dilute xylene solution onto a carbon support film, the structure consists of in-plane meandering cylinders of PE in the PS matrix (Fig. 1a). In order to control the PS-PE microdomain and crystalline PE structures, we first dissolved the diblock in benzoic acid (BA), then crystallized the solution between glass slides in two stages using a modest temperature gradient (10 °C cm^{−1}). In the first directional solidification done at 110 °C, the BA forms large 200 × 500 μm crystals with the (001) planar surfaces parallel to the substrate and the b axis (the fast growth direction) well aligned along the temperature gradient. The remaining solution is then directionally solidified at 60 °C. A thin polymer layer forms adjacent to the glass surfaces, and can be examined by microscopy. A bright field image of the RuO₄-stained film shows well ordered arrays of light, unstained PE domains in a dark, stained PS matrix (Fig. 1b). The light domains are packed on a hexagonal lattice with good long-range order extending over areas of 20 μm diameter.

Tilting the film in the transmission electron microscope (TEM) shows that the light domains are cylindrical rather than spherical. A magnified image shows that the shape of the interface between the PS and PE is non-circular (Fig. 1b inset). The diameter of the PE domains along the b_{BA} direction is about 50% larger on average than in the perpendicular direction (30 nm compared to 20 nm). We will come back to this point later in the context of the dark-field images and the mechanism of structure formation.

An electron diffraction pattern from an unstained film (Fig. 2A) shows very well oriented, almost single-crystal-like, reflections

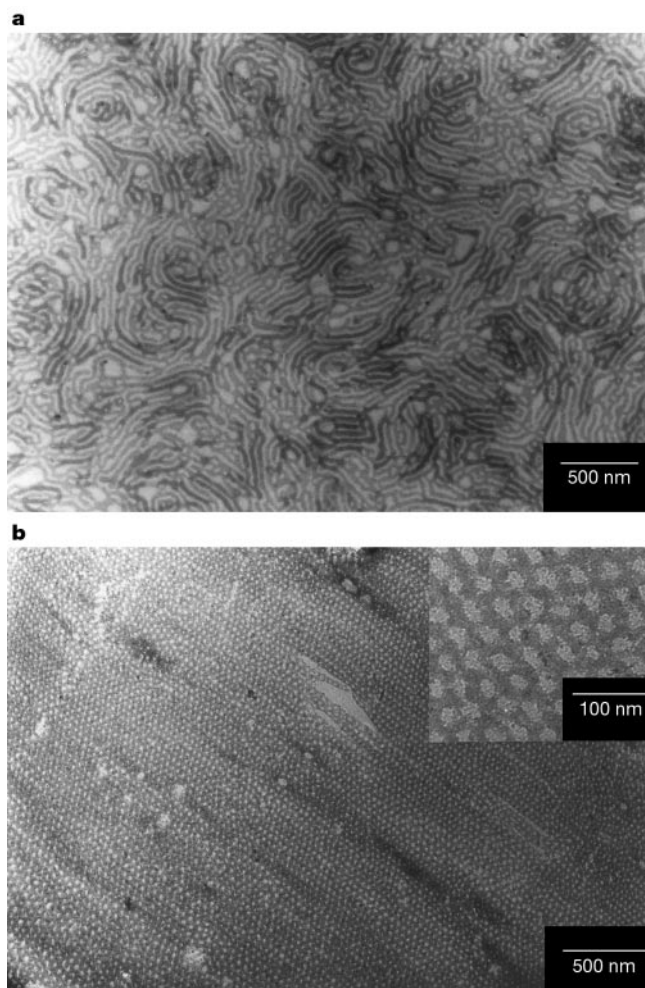


Figure 1 TEM images of solvent-cast, and directionally solidified/epitaxially crystallized, block copolymers. **a**, Solvent-cast thin film of PS-PE block copolymer stained with RuO_4 . The bright-field TEM image shows a poorly ordered microphase-separated structure. The lighter regions correspond to the PE cylindrical domains, the grey regions to the PS matrix. Darker regions appear due to island areas of greater film thickness. **b**, Uniform-thickness, well ordered, thin film of PS-PE block copolymer formed by directional solidification of a

solution of the block copolymer in BA. The PE component forms cylinders oriented perpendicular to the film surface and packed in a pseudo-hexagonal lattice. The interdomain spacing is 40 nm along the b -axis direction of BA, and about 10–15% smaller along the other two directions. Inset, magnified region of **b** showing the non-circular shape of the PS–PE interface.

which index to the $(0kl)$ reciprocal lattice section of orthorhombic polyethylene (space group $Pna2_1$). The (020) reflection of the PE lies along the BA b -axis direction, with the (002) PE chain axis reflection along the perpendicular direction. This situation corresponds to the known epitaxy between homopolyethylene and BA: namely the (100) plane of PE is in contact with the (001) plane of BA, with the b and c axes of PE parallel to the b and a axes of BA, respectively¹⁸. Considering that the sample has a total bulk crystallinity of less than 10%, such a diffraction pattern may be due to the most unusual alignment into the diffraction condition of nearly all of the crystals within the selected area, as well as the enhancement of crystallinity by the substrate²⁰.

The PE crystals can be visualized with dark-field imaging, employing a rotation-tilt stage to bring the strongest-diffracting (110) planes into the Bragg condition (Fig. 2B). Low-dose imaging²¹ shows a periodic array of small rectangular shaped crystals (Fig. 2C). The crystals are arranged in a pseudo-hexagonal lattice whose orientation and average centre-to-centre spacing are in very good correspondence with the orientation and spacings of the cylindrical PE domains seen in the bright-field image. Their size and spacing shows there is precisely one crystalline PE lamella centred in each cylinder, with the b axis of the lamella oriented parallel to the b axis

of the BA, confirming the epitaxy of PE on BA demonstrated by the electron-diffraction pattern. This information also proves that the minority PE component was in direct contact with the BA crystals, and hence is present at least at the BA-facing surface of the thin film. The longest dimension of the PE crystals is along the b axis and parallel to the longest diameter of the non-circular interface of the PE domain. This suggests that during crystallization, the anisotropic growth of the crystallizing PE lamellae deforms the microdomain interface.

The essential aspects of the overall phase transformation can be understood by dividing the process into three stages using a hypothetical solvent–polymer phase diagram (depicted schematically in Fig. 3). An important feature of the phase diagram is the presence of a eutectic due to the intersection of the melting-point depression liquidus curve of the BA with the microphase-separation-transition depression liquidus curve of the block copolymer. The hypoeutectic system begins as a homogeneous liquid (point 1), the temperature is then dropped below the liquidus and growth of BA crystals is initiated by directional solidification (point 2) which increases the copolymer content of the remaining liquid towards the eutectic composition. Further cooling below both the eutectic temperature and the crystallization temperature of the PE results

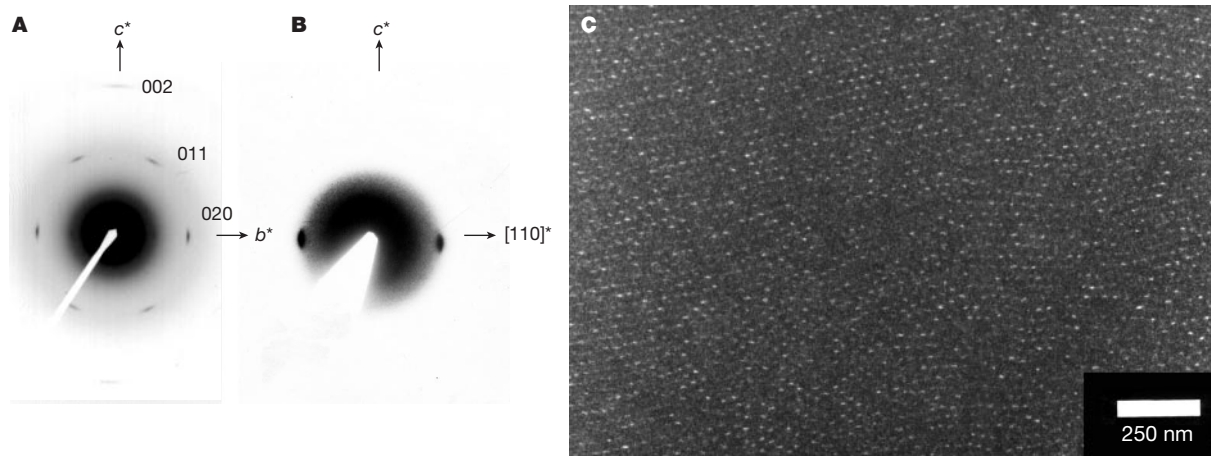


Figure 2 Diffraction patterns and dark-field image of directionally solidified and epitaxially crystallized block copolymer. **A**, Selected-area electron-diffraction pattern of an unstained film of PS-PE block copolymer crystallized from BA, demonstrating epitaxy. The pattern exhibits only $(0k)$ reflections of PE, which indicates that the (100) plane of PE is normal to the electron beam and parallel to the (001) face of the BA crystals. The c and b axes of the PE crystals are parallel to the a and b axes of the BA crystal. **B**, Diffraction pattern from a

similar sample after tilting 34° about the chain-axis direction of PE. The strong (110) diffraction peaks of PE are prominent. **C**, Dark-field image of the PS-PE film using the (110) diffraction spot. Small rectangular PE crystals are observed, well aligned along the b -axis direction of the BA crystals and packed on a pseudo-hexagonal lattice whose size and orientation is the same as seen in Fig. 1b. The PE crystals are 7 nm thick and 20 nm long, with their longest dimension parallel to the $[110]^*$ reciprocal lattice direction.

in the transformation of the eutectic liquid into crystalline BA and an ordered, microphase-separated block copolymer within which the PE block crystallizes (point 3).

The morphological evolution of the system is schematically depicted in Fig. 4a–e. The initial homogeneous solution confined between the glass substrates (Fig. 4a) transforms due to the imposed directional solidification into large crystals of α -BA having (001) surfaces coexisting with a thin liquid layer near the eutectic composition (Fig. 4b). Decreasing the temperature then causes this layer to also directionally solidify by thickening the pre-existing BA crystal and the formation of a thin, metastable vertically oriented lamellar microdomain film (Fig. 4c). This can be understood by noting that in the thin liquid layer, the spatial distribution of the junction points of the block copolymer molecules is nearly random; on rapid extraction of the solvent during the second directional solidification process, the junctions must quickly localize onto periodic interfaces defining the domains. For microdomain interfaces parallel to the fast growth direction, this is readily accomplished, whereas for interfaces that are normal to the fast growth direction, repeated nucleation of these junction surfaces is required. For this 24 vol.% PE diblock, there is more interface per unit volume in a lamellar structure than in a cylindrical one, so the first structure to form is that which forms the fastest—domain layers oriented with their surface normal orthogonal to the fast growth direction of the BA. Of the degenerate set of such orientations, parallel and vertical layers are the most suitable because inclined orientations necessitate junction area density variations near the surface regions.

Depending on the affinity of the respective blocks for the substrate surfaces, the polymer film thickness, microdomain period, and the type of epitaxy and crystallizable block chain orientation with respect to the crystalline substrate, many sequences of events are possible. In the present PS-PE/BA system, for an assumed parallel lamellar orientation, the subsequent transformation into cylinders embedded in a PS matrix would be highly degenerate, so the final morphology would consist of small grains of in-plane cylinders; in fact, the actual structure is well aligned, vertically oriented, pseudo-hexagonally packed PE microdomains with a 10% larger intercylinder spacing along the fast growth direction. For an assumed vertical lamellar domain orientation,

the structure would transform again due to the instability of the flat interface at this composition as well as from the in-plane PE c -axis orientation induced by the epitaxy. The depression of the glass transition temperature of PS by plasticization from the BA (as shown by differential scanning calorimetry, DSC) allows the intermediate structure to reorganize when the crystallization of PE occurs, even at 60°C . Cylinder growth from a vertical layer orientation is more consistent with the observed film morphology, as the domains evolve from an aligned precursor state (Fig. 4d), into the final structure (Fig. 4e).

To check for the stability of the vertically oriented cylindrical structure, the sample was heated to 115°C (above the glass transition temperature of the PS (80°C) and the melting temperature of the PE, but below the melting temperature of BA) for 1 day, and then quenched to vitrify the PS domains and to crystallize the PE block. Note that after about 30 minutes, the BA completely sublimed. TEM showed that the cylinders were still standing up in the thin film, indicating that the vertically oriented structure is a

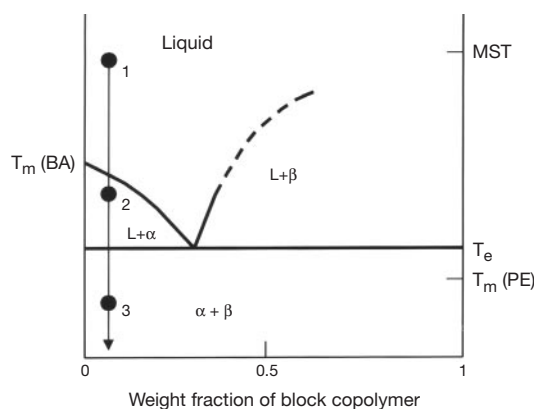


Figure 3 Hypothetical phase diagram of the BA/PS-PE block copolymer system. The figure shows the melting points of benzoic acid ($T_m(\text{BA})$) and polyethylene ($T_m(\text{PE})$), as well as the microphase-separation transition temperature of the PS-PE block copolymer (MST) and the eutectic point and eutectic temperature (T_e). The dilute polymer solution is cooled from location 1 to 3 during the film-formation process.

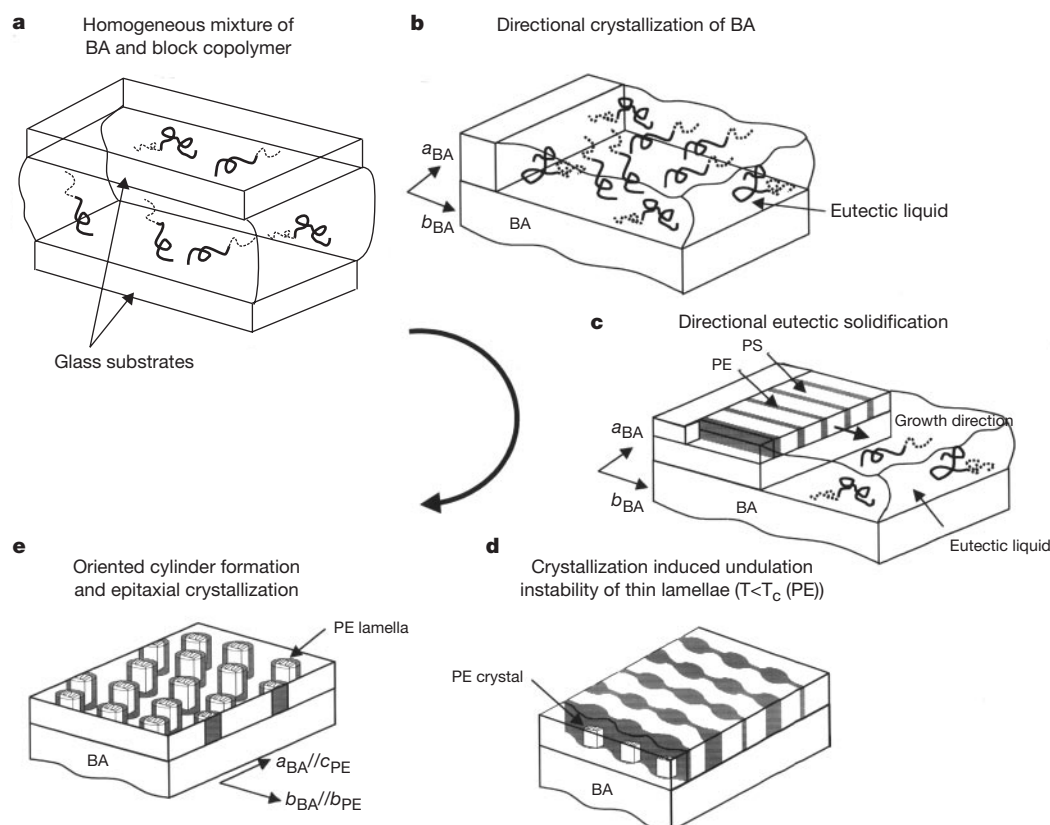


Figure 4 Diagrams showing structural evolution during the directional eutectic solidification and epitaxial crystallization of the block copolymer from the crystallizable solvent. **a**, Homogeneous solution of PS-PE in BA between two glass substrates. **b**, Directional solidification forms crystals of α -BA coexisting with a liquid layer of more concentrated polymer. **c**, Second directional solidification, showing the eutectic liquid layer transforming into BA crystal (which grows on the pre-eutectic α -BA crystal) and an ordered lamellar block copolymer (β). **d**, Because of the highly asymmetric composition of

the block copolymer and the epitaxial crystallization of the PE in contact with the BA substrate, the flat interfaces of vertically oriented lamellae are unstable, and spontaneously deform in order to achieve a more preferred interfacial curvature and allow epitaxial growth of PE. **e**, The layers transform into an array of vertically oriented, pseudo-hexagonally packed semicrystalline PE cylinders. A single, chain-folded PE lamella is formed in each cylinder. The PE crystals have their (100) planes contacting the (001) plane of the BA crystal with $a_{BA} \parallel c_{PE}$ and $b_{BA} \parallel b_{PE}$.

strongly metastable state²². The resultant electron-diffraction pattern of the quenched film showed {110} powder diffraction rings, indicating small, randomly oriented PE crystals.

We have also investigated this and other crystallizable PE block copolymers dissolved in anthracene, and found that the directional solidification and epitaxy process can also readily orient microdomains over large areas in these systems. This suggests that kinetic control of pattern formation via a combination of microphase separation and crystallization along a temperature gradient can organize the block copolymer molecules into ordered microphase-separated structures. This fast process, occurring by the rapid directional extraction of the solvent from the polymer solution by crystallization, may have wide applicability to nanopatterning of thin films. □

Methods

The processing procedure we followed starts with the deposition of a thin (~ 50 – 100 nm) film of the PS-PE copolymer from xylene at room temperature onto a carbon-coated coverslip glass. Crystalline BA powder is then spread over a glass slide, the coverslip is placed polymer-side down on the BA and the system is heated to 150°C , whereupon the BA melts and dissolves the block copolymer. The solution is then supercooled by placing the glass slide on a hot bar at 110°C and touching the edge of the coverslip with tweezers to induce crystallization of the BA (melting temperature $\sim 123^\circ\text{C}$). Rapid crystallization of the BA occurs (growth-front velocity $\sim 2\text{ mm s}^{-1}$) resulting in large, elongated crystals with the b axis parallel to the growth-front direction. Finally the slide is moved to a position on the hot bar with a temperature of 60°C , and held for a minute to complete the crystallization of the BA and the PE, then cooled to room temperature. A razor blade is used to break open the solidified material (cleavage occurs easily on the (001) BA planes). Ethanol at 50°C is then used to wash away the BA from the coverslip. The interior surface

of the coverslip is then scored with a knife, and small sections of the carbon film are floated off onto distilled water and picked up by TEM grids. Some sections were then exposed to RuO_4 to stain the PS regions. The films were examined by TEM in bright and dark field, as well as by selected-area electron diffraction. Polarized light microscopy of the films after BA removal demonstrates that there is a strong texture, most apparent from some elongated regions of thicker polymeric material which formed between the BA crystals. This direction corresponds to the fast growth direction of the BA (the b -axis).

Received 9 July 1999; accepted 12 March 2000.

1. Volkmut, W. D. & Austin, R. H. DNA electrophoresis in microlithographic arrays. *Nature* **358**, 600–602 (1992).
2. Chou, S. Y., Wei, M. S., Krauss, P. R. & Fischer, P. B. Single-domain magnetic pillar array of 35-nm diameter and 65-Gbits/in^2 density for ultra high density quantum magnetic storage. *J. Appl. Phys.* **76**, 6673–6675 (1994).
3. Mansky, P., Chaikin, P. M. & Thomas, E. L. Monolayer films of diblock copolymer microdomains for nanolithographic applications. *J. Mater. Sci.* **30**, 1987–1992 (1995).
4. Park, M., Harrison, C. K., Chaikin, P. M., Register, R. A. & Adamson, D. H. Block copolymer lithography: periodic arrays of $\sim 10^{11}$ holes in 1 square centimeter. *Science* **276**, 1401–1404 (1997).
5. Lammertink, R. G. *et al.* Nanostructured thin films of organic-organometallic block copolymers: one-step lithography with poly(ferrocenylsilanes) by reactive ion etching. *Adv. Mater.* **12**, 98–103 (2000).
6. Morkved, T. L. *et al.* Local control of microdomain orientation in diblock copolymer thin films with electric fields. *Science* **273**, 931–933 (1996).
7. Mansky, P. *et al.* Large-area domain alignment in block copolymer thin films using electric fields. *Macromolecules* **31**, 4399–4401 (1998).
8. Fasolka, M. *et al.* Observed substrate topography-mediated lateral patterning of diblock copolymer films. *Phys. Rev. Lett.* **79**, 3018–3021 (1997).
9. Huang, E., Rockford, L., Russell, T. P. & Hawker, C. J. Nanodomain control in copolymer thin films. *Nature* **395**, 757–758 (1998).
10. Muthukumar, M., Ober, C. K. & Thomas, E. L. Competing interactions and levels of ordering in self-organizing polymeric materials. *Science* **277**, 1225–1232 (1997).
11. Hashimoto, T., Shibayama, M., Fujimura, M. & Kawai, H. in *Block Copolymers, Science and Technology* (ed. Meier, D. J.) 63–108 (Harwood Academic, London, 1983).
12. Bodcomb, J., Funaki, Y., Kimishima, K. & Hashimoto, T. Single grain lamellar microdomain from a diblock copolymer. *Macromolecules* **32**, 2075–2077 (1999).

13. Flemings, M. C. *Solidification Processing* (McGraw Hill, New York, 1974).
14. Swei, G. S., Lando, J. B., Reichert, S. E. & Mauritz, K. A. *Encyclopedia of Polymer Science and Engineering* Vol. 6, p. 209 (Wiley, New York, 1986).
15. Wittmann, J. C. & Lotz, B. Epitaxial crystallization of polymers on organic and polymeric substrates. *Prog. Polym. Sci.* **15**, 909–948 (1990).
16. Reiter, G. *et al.* Nanometer scale surface patterns with long range order created by crystallization of diblock copolymers. *Phys. Rev. Lett.* **83**, 3844–3847 (1999).
17. Smith, P. & Pennings, A. J. Eutectic crystallization of pseudo binary systems of polyethylene and high melting diluents. *Polymer* **15**, 413–419 (1974).
18. Wittmann, J. C., Hodge, A. M. & Lotz, B. Epitaxial crystallization of polymers onto benzoic-acid-polyethylene and paraffins, aliphatic polyesters, and polyamides. *J. Polym. Sci. Polym. Phys. Edn* **21**, 2495–2509 (1983).
19. Dorset, D. L., Hanlon, J. & Karet, G. Epitaxy and structure of paraffin-diluent eutectics. *Macromolecules* **22**, 2169–2176 (1989).
20. Wu, S. *Polymer Interface and Adhesion* (Marcel Dekker, New York, 1982).
21. Thomas, E. L. & Ast, D. Image intensification and the electron microscopy of radiation sensitive polymers. *Polymer* **15**, 37–41 (1974).
22. Kim, G. & Libera, M. Kinetic constraints on the development of surface microstructures in SBS thin films. *Macromolecules* **31**, 2670–2672 (1998).

Acknowledgements

We thank J.C. Wittmann, Y. Cohen, C. Thompson, C. Carter and L. Fetters for conversations; L. Fetters also synthesized the copolymer. This work was supported by the NSF, ACS-PRF and US-France NSF-CNRS.

Correspondence and requests for materials should be addressed to E.L.T. (e-mail: elt@mit.edu).

Large-scale synthesis of a silicon photonic crystal with a complete three-dimensional bandgap near 1.5 micrometres

Alvaro Blanco^{*†‡}, Emmanuel Chomski[§], Serguei Grabtchak^{*}, Marta Ibasate^{†‡}, Sajeev John^{*}, Stephen W. Leonard^{*}, Cefe Lopez^{†‡}, Francisco Meseguer^{†‡}, Hernan Miguez^{†‡}, Jessica P. Mondia^{*}, Geoffrey A. Ozin[§], Ovidiu Toader^{*} & Henry M. van Driel^{*}

^{*} Department of Physics, 60 Saint George Street, University of Toronto, Toronto, Ontario, Canada M5S 1A7

[†] Unidad Asociada (CSIC-UPV) Universidad Politécnica, 46022 Valencia, Spain

[‡] Instituto de Ciencia de Materiales de Madrid (CSIC), Cantoblanco 28049, Madrid, Spain

[§] Department of Chemistry, 80 Saint George Street, University of Toronto, Toronto, Ontario, Canada M5S 3H6

Photonic technology, using light instead of electrons as the information carrier, is increasingly replacing electronics in communication and information management systems. Microscopic light manipulation, for this purpose, is achievable through photonic bandgap materials^{1,2}, a special class of photonic crystals in which three-dimensional, periodic dielectric constant variations controllably prohibit electromagnetic propagation throughout a specified frequency band. This can result in the localization of photons^{3–6}, thus providing a mechanism for controlling and inhibiting spontaneous light emission that can be exploited for photonic device fabrication. In fact, carefully engineered line defects could act as waveguides connecting photonic devices in all-optical microchips⁷, and infiltration of the photonic material with suitable liquid crystals might produce photonic bandgap structures (and hence light-flow patterns) fully tunable by an externally applied voltage^{8–10}. However, the realization of this technology requires a strategy for the efficient synthesis of high-quality, large-scale photonic crystals with photonic bandgaps at

micrometre and sub-micrometre wavelengths, and with rationally designed line and point defects for optical circuitry. Here we describe single crystals of silicon inverse opal with a complete three-dimensional photonic bandgap centred on 1.46 μm , produced by growing silicon inside the voids of an opal template of close-packed silica spheres that are connected by small 'necks' formed during sintering, followed by removal of the silica template. The synthesis method is simple and inexpensive, yielding photonic crystals of pure silicon that are easily integrated with existing silicon-based microelectronics.

Previously, three-dimensional photonic crystals have been fabricated in silicon by complex lithographic procedures which resulted in a thickness of only two unit cells in the growth direction^{11,12}. Our materials synthesis, in contrast, begins with the highly controlled formation of a weakly sintered f.c.c. lattice of mono-disperse silica spheres whose diameter can be chosen between 600 and 1,000 nm. The silica opal structure is then used as a template for silicon infiltration. The large silica sphere size ensures that the final structure will have a complete photonic bandgap (PBG) in a wavelength range above the optical absorption edge ($\sim 1.1 \mu\text{m}$) of bulk silicon. The synthesis of smaller (0.2–0.6 μm) sphere opals has been described previously¹³. The fabrication of periodically ordered opals with the larger unit cells involves the synthesis of mono-disperse ($< 5\%$ diameter variation) large spheres by a modified Stöber method¹⁴. The synthesis is performed in a suspension of smaller (500 nm diameter), mono-disperse silica particles that act as seeds for further growth; this leads to a homogeneous increase of the sphere size. Once the large spheres are made, they are settled in an aqueous solution of ethylene glycol (typically 50% concentration). The resulting opal is a close-packed f.c.c. lattice of silica spheres with a typical single domain size of 100 μm . A further sintering process leads to the formation of small necks between the silica spheres and

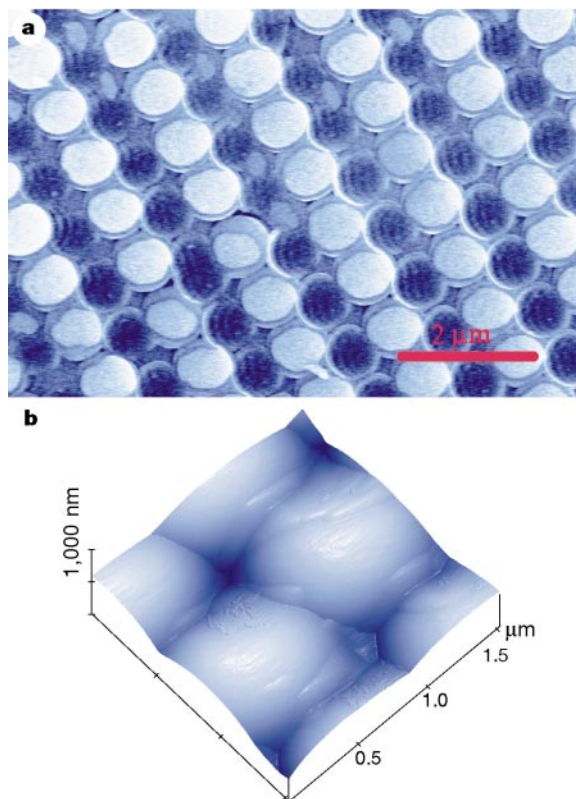


Figure 1 Silicon-infiltrated opal. **a**, SEM image of an internal [113] face of the silicon-infiltrated opal. The silica is indicated by light blue. **b**, AFM image of a 1.5 μm by 1.5 μm area of the infiltrated opal surface, highlighting the smoothness of the silicon coating.

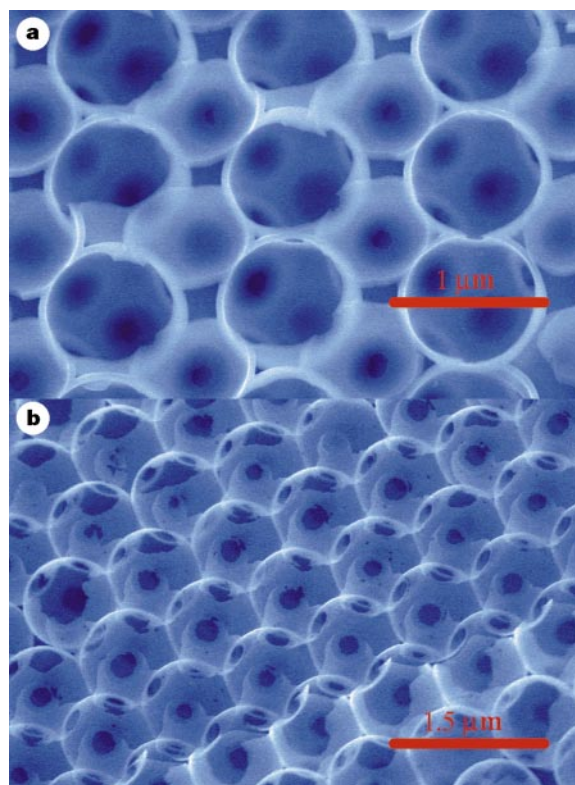


Figure 2 SEM images of internal facets of silicon inverse opal: **a**, [110] facet. **b**, [111] facet.

provides the template with mechanical stability. The necks also help to control the opal void volume for subsequent synthesis¹⁵, and induce a connected network topology, allowing for removal of the template by an acid etching. This network topology also enhances optical scattering effects that are important for PBG formation¹⁶.

Silicon is grown inside the voids of the opal template by means of chemical vapour deposition (CVD) using disilane (Si_2H_6) gas as a precursor^{17,18}. Silane and disilane gas have both been used to grow silicon nanoclusters in periodic micro- and mesoporous silicas^{7–19}. In our experiment, the deposition temperature is varied from 250 °C for low filling fractions to 350 °C for high ones. Our band structure calculations predict that the maximum PBG width is obtained with a 90–97% filling of the opal voids in the form of a uniformly thick wetting layer on the silica surfaces¹⁶. Samples with degrees of infiltration ranging up to 100% were synthesized. The reaction time was typically 24 hours and the disilane pressure was about 200 torr. After silicon growth, the samples are heated to 600 °C in order to improve the semiconductor crystallization and to allow diffusion of silicon inside the void structure. The silica template is subsequently removed using a fluoride-based etching procedure designed to minimize the dissolution of the macroporous silicon backbone.

Microscope Raman spectroscopy was used to ascertain the sample quality. A single phonon peak was observed in our samples centred at 519 cm^{-1} with a width less than 8 cm^{-1} consistent with the presence of crystalline silicon. Typical scanning electron microscopy (SEM) and atomic force microscopy (AFM) images of the infiltrated silica opal are shown in Fig. 1. The SEM image indicates a thick, uniform layer of silicon surrounding the well-ordered silica spheres (shown in light blue), indicating a high degree of infiltration. From the AFM image the silicon surface roughness is estimated to be less than 2 nm. The growth of the silicon-wetting layer is quite homogeneous and is independent of the local characteristics of the opal template. A SEM image of a typical inverse silicon opal taken after etching the infiltrated structure is shown in Fig. 2. The images reveal

an interconnected network of air spheres surrounded by silicon shells, inheriting the f.c.c. order of the opal template. Adjacent air spheres are connected by windows, which result from the sintering process.

Unlike earlier inverse opal structures made of TiO_2 (refs 20, 21), graphitic carbon²², CdSe and CdS (refs 23, 24), our silicon inverse opal structure simultaneously satisfies the two essential criteria for complete PBG formation. First, the refractive index of silicon (3.45 at $1.5\text{ }\mu\text{m}$) is well above the theoretically determined¹⁶ threshold (2.8) for a PBG in a f.c.c. lattice of air spheres. Second, the optical absorption edge of the silicon backbone occurs at a wavelength well below the PBG, thereby allowing coherent localization of light within the material with minimal absorptive losses. Such localization is an essential feature for PBG device applications; as light is caged within the dielectric microstructure, it cannot scatter into unwanted modes of free propagation and is forced to flow along engineered defect channels between devices integrated into an optical microchip.

The theoretically predicted photonic band structure of a silicon inverse opal, with 88% infiltration of silicon into the available volume of the opal template voids, is shown in Fig. 3. The calculations were performed using the plane-wave expansion method with a basis of 2662 plane waves²⁵. As the infiltration level is numerically increased from 88% to 100%, the centre of the complete PBG moves from 1.46 to $1.55\text{ }\mu\text{m}$, spanning the wavelength range of choice for fibre optic telecommunications.

The reflection and transmission spectra of a silicon inverse opal were measured using a Bomem Fourier transform infrared spectrometer. A microscope attached to the spectrometer was used to illuminate a single crystal domain with a high surface quality. Unpolarized light passed through the microscope produced a spot size $\sim 40\text{ }\mu\text{m}$ in diameter with the incident light cone spanning 15° – 35° from normal incidence. The resulting normalized reflection and transmission spectra are shown in Fig. 4. The reflection spectrum exhibits a broad peak with a centre wavelength of $2.5\text{ }\mu\text{m}$ with three

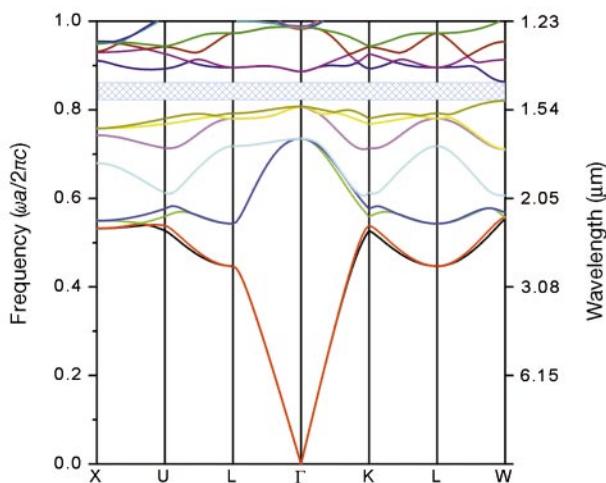


Figure 3 Band structure of silicon inverse opal with an 88% infiltration of Si into the available opal template voids. The complete photonic bandgap is shown by the crosshatched region, with a gap to mid-gap ratio of 5.1%.

additional peaks between 1.2 and 1.8 μm . The reflectivity peak amplitude exceeds 80% and 70% at 2.5 μm and 1.46 μm , respectively, indicating good crystal quality. The sample transmission spectrum, indicated in the inset, has been corrected for diffuse scattering²⁶. We note a strong correlation between the key spectral features of the transmission and reflectance data, that is, the maxima in the reflectance spectrum correspond to the minima in the transmission spectrum.

To compare the optical measurements to band structure calculations we first determined the lattice constant. This is preserved after infiltration and inversion, and was determined from reflectivity measurements of the bare opal at normal incidence. By comparing the spectral position of the first stop band edges (associated with the pseudogap at the L point of the first Brillouin zone²⁷) to those predicted by band structure calculations (using a silica refractive index of 1.45), we obtained a (cubic) lattice constant of 1.23 μm .

The degree of infiltration was obtained by varying its value in the band structure calculation to obtain agreement with the centre wavelength of the first stop band, measured to be 2.5 μm . The silicon shell topology of Fig. 2 was assumed. One sees from Fig. 4 that the calculation gives the appropriate width for the stop band. Excellent agreement is also observed between the position of the peak in the reflectivity spectrum and the centre of the stop band at 1.46 μm , which is highly sensitive to the degree of infiltration. Calculations identify this band as the frequency interval associated with a complete bandgap. The good agreement between the measured spectra and the calculated band structure suggests the presence of a complete PBG centred at 1.46 μm , with a gap to mid-gap ratio of 5.1%. More comprehensive measurements are needed to provide full experimental verification of the photonic bandgap.

The synthesis of a very-large-scale, silicon-based PBG material offers a number of imminent possibilities, involving further infiltration of this highly open structure with light-emitting molecules or atoms. The resulting luminescence and lasing characteristics of light emitting species near a three-dimensional photonic band edge are expected to be quite striking. Important low-threshold all-optical switching effects and an anomalous nonlinear optical response have been predicted²⁵. In this regard, it is useful to explore self-assembly methods for creating diamond lattice templates from which a considerably larger PBG may be achieved²⁸. It is also important to generalize the template-formation procedure in order to engineer wave-guide channels and specified point defects through which light can flow. Methods of soft lithography and

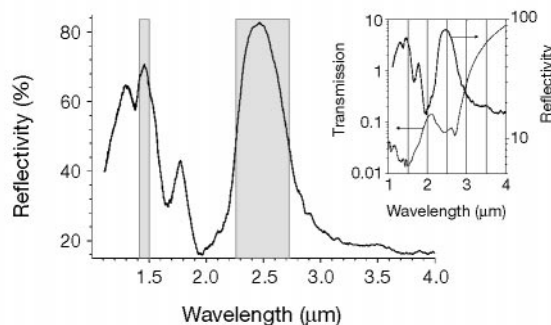


Figure 4 Reflection spectrum from the inverse silicon opal. Shaded regions at 2.5 μm and 1.5 μm show the calculated positions of the first stop band and the complete photonic bandgap, respectively. Inset, the transmission spectrum of the same sample on a logarithmic scale. The narrow transmission minimum at 2.7 μm is due to water absorption.

micromoulding in capillaries²⁹ may soon prove effective in realizing such 'circuits of light'.

Received 31 January; accepted 14 April 2000.

- John, S. Strong localization of photons in certain disordered dielectric lattices. *Phys. Rev. Lett.* **58**, 2486–2489 (1987).
- Yablonovitch, E. Inhibited spontaneous emission in solid-state physics and electronics. *Phys. Rev. Lett.* **58**, 2059–2062 (1987).
- John, S. Electromagnetic absorption in a disordered medium near a photon mobility edge. *Phys. Rev. Lett.* **53**, 2169–2172 (1984).
- Anderson, P. W. The question of classical localization: A theory of white paint? *Phil. Mag. B* **52**, 505–509 (1985).
- John, S. Localization of light. *Phys. Today* **44**, 32–40 (1991).
- Wiersma, D. S., Bartolini, P., Lagendijk, A. & Righini, R. Localization of light in a disordered medium. *Nature* **390**, 671–673 (1997).
- Joannopoulos, J. D., Villeneuve, P. R. & Fan, S. Photonic crystals: putting a new twist on light. *Nature* **386**, 143–149 (1997).
- Busch, K. & John, S. Liquid-crystal photonic-band-gap materials: The tunable electromagnetic vacuum. *Phys. Rev. Lett.* **83**, 967–970 (1999).
- Yablonovitch, E. Liquid versus photonic crystals. *Nature* **401**, 539–541 (1999).
- Leonard, S. W. et al. Tunable two-dimensional photonic crystals using liquid crystal infiltration. *Phys. Rev. B* **61**, 2389–2392 (2000).
- Lin, S. Y. & Fleming, J. G. A. Three-dimensional optical photonic crystal. *IEEE J. Lightwave Technol.* **17**, 1944–1947 (1999).
- Noda, S., Yamamoto, N., Imada, M., Kobayashi, H. & Okano, M. A three-dimensional optical photonic crystal. *IEEE J. Lightwave Technol.* **17**, 1948–1955 (1999).
- Míguez, H. et al. Photonic crystal properties of packed submicrometric SiO_2 spheres. *Appl. Phys. Lett.* **71**, 1148–1150 (1997).
- Stöber, W., Fink, A. & Bohn, E. Controlled growth of monodisperse silica spheres in the micron size range. *J. Colloid Interface Sci.* **26**, 62–69 (1968).
- Míguez, H. et al. Control of the photonic crystal properties of fcc packed submicrometer SiO_2 spheres by sintering. *Adv. Mater.* **10**, 480–483 (1999).
- Busch, K. & John, S. Photonic band gap formation in certain self-organizing systems. *Phys. Rev. E* **58**, 3896–3908 (1998).
- Chomski, E., Dag, Ö., Kuperman, A. & Ozin, G. A. New forms of luminescent silicon: silicon-silica composite mesostructures. *Chem. Vap. Dep.* **2**, 8–13 (1996).
- Dag, Ö., Ozin, G. A., Yang, H., Reber, C. & Bussiére, G. Photoluminescent silicon clusters in oriented hexagonal mesoporous silica film. *Adv. Mater.* **11**, 474–480 (1999).
- Bogomolov, N. N. et al. Fabrication of regular three-dimensional lattices of submicron silicon clusters in a SiO_2 matrix. *Pis'ma Zh. Tekh. Fiz.* **24**, 90–95 (1998); also *Technol. Phys. Lett.* **24**, 326–327 (1998).
- Holland, B. T., Blanford, C. F. & Stein, A. Synthesis of macroporous minerals with highly ordered three-dimensional arrays of spheroidal voids. *Science* **281**, 538–540 (1998).
- Wijnhoven, J. E. G. J. & Vos, W. L. Preparation of photonic crystals made of air spheres in titania. *Science* **281**, 802–804 (1998).
- Zakhidov, A. A. et al. Carbon structures with three-dimensional periodicity at optical wavelengths. *Science* **282**, 897–901 (1998).
- Vlasov, Y. A., Yao, N. & Norris, D. J. Synthesis of photonic crystals for optical wavelengths from semiconductor quantum dots. *Adv. Mater.* **11**, 165–169 (1999).
- Braun, P. V. & Wiltzius, P. Electrochemically grown photonic crystals. *Nature* **402**, 603–604 (1999).
- John, S. & Busch, K. Photonic bandgap formation and tunability in certain self-organizing systems. *J. Lightwave Technol.* **17**, 1931–1943 (1999).
- Vlasov, Y. A., Astratov, V. N., Karimov, O. Z. & Kaplyanskii, A. A. Existence of a photonic pseudogap for visible light in synthetic opals. *Phys. Rev. B* **55**, R13357–R13360 (1997).
- Ashcroft, N. & Mermin, D. *Solid State Physics* 161 (Holt, Rinehart & Winston, New York, 1976).
- Ho, K. M., Chan, C. T. & Soukoulis, C. M. Existence of a photonic gap in periodic dielectric structures. *Phys. Rev. Lett.* **65**, 3152–3155 (1990).
- Xia, Y. & Whitesides, G. M. Soft lithography. *Angew. Chem. Int. Edn Engl.* **37**, 550–575 (1998).

Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, Photonics Research Ontario, the Fundacion Ramon Areces, the Spanish CICYT project and the European Community Project. We are grateful to I. Sokolov and N. Coombs for their technical assistance with the AFM and SEM images (Figs 1 and 2).

Correspondence and requests for materials should be addressed to S. J. (e-mail: john@physics.utoronto.ca).

Deep convective clouds with sustained supercooled liquid water down to -37.5°C

Daniel Rosenfeld* & William L. Woodley†

* Institute of Earth Sciences, the Hebrew University of Jerusalem, Jerusalem 91904, Israel

† Woodley Weather Consultants, 11 White Fir Court, Littleton, Colorado 80127, USA

In cirrus¹ and orographic wave clouds², highly supercooled water has been observed in small quantities (less than 0.15 g m^{-3}). This high degree of supercooling was attributed to the small droplet size and the lack of ice nuclei at the heights of these clouds^{1,2}. For deep convective clouds, which have much larger droplets near their tops and which take in aerosols from near the ground, no such measurements have hitherto been reported. However, satellite data suggest that highly supercooled water (down to -38°C) frequently occurs in vigorous continental convective storms³. Here we report *in situ* measurements in deep convective clouds from an aircraft, showing that most of the condensed water remains liquid down to -37.5°C . The droplets reach a median volume diameter of $17\text{ }\mu\text{m}$ and amount to 1.8 g m^{-3} , one order of magnitude more than previously reported². At slightly colder temperatures only ice was found, suggesting homogeneous freezing. Because of the poor knowledge of mixed-phase cloud processes⁴, the simulation of clouds using numerical models is difficult at present. Our observations will help to understand these cloud processes, such as rainfall, hail, and cloud electrification, together with their implications for the climate system.

Cloud droplets do not readily freeze at 0°C , but often remain liquid at colder temperatures in a "supercooled" state. Freezing of the cloud droplets can then be triggered by ice nuclei, or the droplets can freeze in an homogeneous fashion. The lowest temperatures at which pure water droplets can exist in a supercooled state for times longer than a fraction of a second before homogeneously freezing depends on the drop size. According to both theory⁵ and laboratory experiments⁶, a $10\text{-}\mu\text{m}$ cloud droplet freezes homogeneously near -39°C . The coldest previously reported *in situ* measurements of supercooled liquid water content (SLWC) at temperatures below -32°C , in excess of the sensitivity of the measuring instruments (0.02 g m^{-3} for hot-wire probes⁷), was 0.14 g m^{-3} at -36°C in orographic lenticular wave clouds². An SLWC of 0.06 g m^{-3} at the base of cirrus clouds, between -35° and -36°C , was reported by Sassen¹, who stated that "in comparison with earlier reported aircraft measurements, the detection of such highly supercooled water is unique".

Both reports^{1,2} suggested that the dearth of ice nuclei derived from the Earth's surface in the upper troposphere prevented heterogeneous nucleation, allowing for the observed homogeneous nucleation at such cold temperatures in altocumulus and cirrus

clouds⁸. No previous reports are available for observations of similar highly supercooled water and homogeneous freezing in convective clouds with roots near the surface. In view of the reported findings to the contrary here, we can only speculate about the reasons they have not been reported previously. Perhaps low priority was given to such measurements, because it was felt that clouds taking in air rich in ice nuclei from the boundary layer would glaciate long before reaching the point of homogeneous freezing⁹. Or perhaps such measurements were not made because of the safety problems involved in penetrating vigorous cumulonimbus towers at the -30 to -40°C isotherm levels in storms that typically contain hail and frequent lightning.

The first indications available to us that highly supercooled water might exist in convective clouds were obtained by remote sensing from satellites over Thailand, using the technique reported in ref. 3. The inference of supercooled water at temperatures below -30°C prompted us to fly with the Thai King Air cloud-physics aircraft to measure the cloud microstructure in Thailand clouds. Penetrating feeders of cumulonimbus clouds, an SLWC of 2.4 g m^{-3} was measured at the operational ceiling of the aircraft ($9,300\text{ m}$ above sea level, a.s.l.) at a temperature of -31.6°C (ref. 10). The cloud-base temperature was $+13^{\circ}\text{C}$ at $2,800\text{ m a.s.l.}$, and the SLWC was measured by the King hot-wire instrument.

Analysis of the satellite data using the methodology of Rosenfeld and Lensky³ suggested that supercooled water occasionally occurred at temperatures approaching -40°C in cumulonimbus clouds over the western USA. An opportunity to validate these satellite inferences came when Weather Modification Inc. gave us access to its Lear jet cloud-physics aircraft in the period 9–14 August 1999 for measurements in Texas clouds. The aircraft cloud-physics instrumentation included the following: (1) a hot-wire cloud liquid water probe, model LWC-100 (Droplet Measurement Technologies). The measurement efficiency of the sensor depends on the drop sizes. For the observed distribution, it underestimated the SLWC by not more than 10%. (2) An air temperature probe, model 102AU1AP (Rosemount). (3) A forward scattering spectrometer probe (FSSP) for droplet sizes in the range $0.5\text{--}47\text{ }\mu\text{m}$, model FSSP-100 (Particle Measuring Systems Inc.). (4) An optical array particle imaging probe, for the range $25\text{--}800\text{ }\mu\text{m}$, model OAP-2D2-C (Particle Measuring Systems, Inc.).

High-altitude measurements in the tops of vigorous growing convective elements of cumulonimbus clouds were done with the Lear jet on 11 and 13 August 1999. The visibly most-vigorous new convective elements were measured as they grew through the measurement flight level. In addition, extensive measurements were made at all levels down to the cloud base, in order to document the vertical microphysical evolution of the cloud. Cloud base on both days was near $3,500\text{ m a.s.l.}$ at a temperature of 10°C .

On the flight of 11 August (20:28–23:42 GMT) just to the west of Lubbock, Texas (34°N , 102°W), an SLWC of 0.6 g m^{-3} was observed at -35.9°C , 0.9 g m^{-3} at -35.6°C , and 1.5 g m^{-3} at -34.4°C . Larger values of SLWC, up to 2.4 g m^{-3} , were recorded at warmer cloud temperatures. Direct measurements of the updraught velocity were not available. However, a rate of climb of 6 m s^{-1} was sufficient to keep up with the rate of growth of the tops of some of the clouds containing highly supercooled water. The residence time of the water was measured by repeated penetrations in the same cloud, which was narrow and clearly isolated. In three passes spaced at 3.5-min intervals, the temperatures and maximum cloud water contents were 1.2 g m^{-3} at -32.9°C , 1.5 g m^{-3} at -32.7°C , and 0.4 g m^{-3} at -35.2°C . This means that the large amounts of highly supercooled cloud water were not a transient feature, but rather were long lasting, with freezing time of about 7 min.

On the flight of 13 August (20:26–23:09 GMT), to the north of Midland, Texas (33°N 102°W), effort was focused on documentation of the transition from water clouds to ice clouds in vigorous convective elements. Extensive documentation of the clouds from

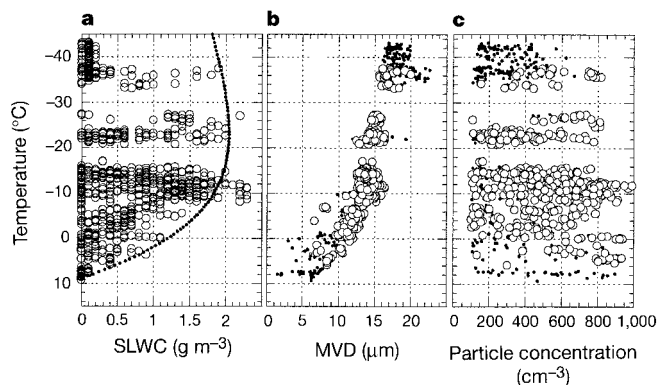


Figure 1 Cloud microstructure and composition as a function of temperature. Data were obtained on 13 August 1999. Each point represents one second of measurements, or 150–200 m of cloud path. **a**, The supercooled liquid water content (SLWC). Note the abrupt decrease of SLWC at -38°C , indicating the point of homogeneous freezing. The dotted curve marks half the adiabatic value of the water content, as a reference for the rate of consumption of cloud water by freezing and other processes. **b**, The median volume diameter (MVD; as measured by the FSSP instrument) of the cloud particles for all points containing more than 100 particles cm^{-3} . The open circles denote cloud segments with liquid water contents greater than 0.3 g m^{-3} , as measured by the hot-wire instrument. **c**, As **b**, but for the FSSP-measured concentration of cloud particles.

their bases was done as on the 11 August. The vertical evolution of cloud microstructure was found to be similar for both days. We therefore show in Fig. 1 only data from 13 August. As illustrated in Fig. 1a, maximum SLWC values of nearly half the adiabatic water content was measured throughout the cloud depth, up to the -37.5°C isotherm. The aircraft windscreens were instantly covered with rime ice during readings of high SLWC despite the windscreens heater, showing that there really was ample supercooled water at these very low temperatures. An abrupt vanishing of the SLWC was recorded at colder temperatures in the same clouds. The remaining reading of up to 0.2 g m^{-3} in obviously fully glaciated clouds (that is, at temperatures $< -40^{\circ}\text{C}$, where theory dictates that droplets greater than $1 \mu\text{m}$ in size must freeze homogeneously⁵) is attributed to the cooling effect of the frozen cloud drops on the hot wire (D. Baumgardner, personal communication). The lack of significant freezing at higher temperatures, and the sudden freezing at -38°C , indicates that homogeneous freezing is the main glaciating mechanism in the measured convective clouds. This observation has a number of implications.

First, heterogeneous nucleation by ice nuclei was incapable of freezing much of the cloud water. However, it is not likely that this was because of a dearth of ice nuclei. The air ingested into the cloud base had substantial amounts of aerosols of unknown composition. The FSSP measured average concentrations of $>0.5\text{-}\mu\text{m}$ particles of $0.5 \text{ particles cm}^{-3}$. Most insoluble aerosols of that size become ice nuclei below -25°C . A possible explanation might be the very small collision efficiencies between the $<200\text{-}\mu\text{m}$ ice crystals that were nucleated heterogeneously and the $<40\text{-}\mu\text{m}$ cloud droplets (ref. 4).

Second, most of the water remained in the form of large concentrations of small cloud droplets (Fig. 1b), with FSSP-measured median volume diameter (MVD) increasing with height, reaching $17 \mu\text{m}$ at the -37.5°C isotherm, just below the height of homogeneous freezing (see Fig. 1c).

Third, although the FSSP cannot resolve ice from water, the fact that the SLWC was half the adiabatic water content implies the existence of liquid water droplets of at least the MVD size of $17 \mu\text{m}$. Such droplets freeze homogeneously near -38°C (ref. 5).

Fourth, the nearly constant water content (of half the adiabatic value; see Fig. 1b) and the small MVD of the cloud droplets suggest

that only a small fraction of the cloud water was converted into precipitation. That is confirmed by the optical array particle imaging probe measurements, showing that ice particles of greater than $\sim 50 \mu\text{m}$ exceed a concentration of $\sim 10 \text{ l}^{-1}$ in the high-SLWC clouds only at temperatures below -30°C .

Fifth, the maintenance of the MVD through the glaciation at -38°C indicates that the cloud droplets freeze into small ice particles, presumably in the form of frozen droplets. This ice leaves the cloud through the cumulonimbus anvil, and is not likely to contribute to the precipitation. Therefore, the observation of apparent homogeneous freezing in cumulonimbus clouds indicates very poor precipitation efficiency.

Sixth, the deep layer of high SLWC provides large growth potential for the small concentration of ice precipitation particles that are initiated in the lower parts of the updraft. This means that such highly supercooled and persistent SLWC represents favourable conditions for the growth of large hail.

Last, given the empirical necessity for liquid water for appreciable electrification in laboratory experiments¹¹, the extension of supercooled droplets to greater heights above the 0°C isotherm might explain observations¹² that deep continental clouds have the strongest electrification of all cumulonimbus types.

The accuracy of the instrument reading is critical in the context of these findings. A calibration check was made to the instruments before and after the reported flights, and they were found to have been properly calibrated. The temperature readings were validated against the radiosonde balloon, which was launched from Midland at 00 GMT on 14 August. The sounding was representative, because it was within 100 km and 1 h of the time of the aircraft measurements. The comparison of the aircraft-measured temperature (T_a) with the sounding temperature (T_s) at the point where 1.8 g m^{-3} of SLWC was measured shows the following. In cloud, at a static pressure of 265 hPa, $T_a = -37.5^{\circ}\text{C}$; T_s at the same pressure was -38.1°C . Out of cloud in level flight, $T_a = -37.5^{\circ}\text{C}$ at a static pressure of 260 hPa, whereas T_s was -39.3°C at the same pressure. Direct checks of the temperature probe on the ground provided: $T_a = +0.4^{\circ}\text{C}$ at 0°C , and $T_a = -9.7^{\circ}\text{C}$ at -10°C . These comparisons suggest that our measurements provide a conservative estimate of the temperature at which homogeneous freezing took place. It might in fact have occurred at a slightly lower temperature.

We note that the Lear jet was available for only 6 days, and on only 2 days were attempts made to document the highly supercooled portions of the clouds, although potentially suitable clouds occurred also on other days of that week. The fact that the highly supercooled water was found in both the two attempts strengthens the evidence from satellite indications that the reported observations are not rare occurrences.

We report here that, in some clouds, most of the condensed water remains liquid until the point of homogeneous freezing; we consider that this will necessitate a significant revision of cloud models simulating rain, hail, and cloud electrification. Incorporation of these changes in global circulation models will probably produce substantial changes in our understanding of the way that aerosols, clouds and precipitation are affecting the global climate. For example, aerosols that serve as small cloud condensation nuclei—which make the clouds more continental (that is, with smaller droplets)—might reduce their precipitation efficiency and thus inhibit rainfall in polluted areas. Reduced precipitation would mean more water vapour flux to the upper troposphere and lower stratosphere, and reduced net latent heat release, which in fact would change both the radiative and heating forcing of the climate system.

Note added in proof: Observations conducted by the authors in the last week of January and the first week of February 2000 at Mendoza, Argentina, documented five more cases of supercooled cloud water: up to 4 g m^{-3} of supercooled cloud water was found at -38°C , vanishing abruptly at lower temperatures. □

Received 3 September 1999; accepted 17 April 2000.

1. Sassen, K. Highly supercooled cirrus cloud water: confirmation and climatic implications. *Science* **227**, 411–413 (1985).
2. Heymsfield, J. A. & Miloshevich, L. M. Homogeneous ice nucleation and supercooled liquid water in orographic wave clouds. *J. Atmos. Sci.* **50**, 2335–2353 (1993).
3. Rosenfeld, D. & Lensky, I. M. Spaceborne sensed insights into precipitation formation processes in continental and maritime clouds. *Bull. Am. Meteorol. Soc.* **79**, 2457–2476 (1998).
4. Pruppacher, H. R. & Klett, J. D. *Microphysics of Clouds and Precipitation* 2nd edn (Kluwer Academic, Norwell, Massachusetts, 1997).
5. Jeffery, C. A. & Austin, P. H. Homogeneous nucleation of supercooled water: results from a new equation of state. *J. Geophys. Res.* **102**, 25269–25279 (1997).
6. Pruppacher, H. R. A new look at homogeneous ice nucleation in supercooled water drops. *J. Atmos. Sci.* **52**, 1924–1933 (1995).
7. Heymsfield, J. A. & Miloshevich, L. M. Evaluation of liquid water measuring instruments in cold clouds, sampled during FIRE. *J. Atmos. Ocean. Technol.* **6**, 378–388 (1989).
8. Sassen, K. in *Nucleation and Atmospheric Aerosols* (eds Fukuta, N. & Wagner, P. E.) 287–290 (Deepak, Hampton, Virginia, 1992).
9. Houghton, H. G. *Physical Meteorology* (MIT Press, Cambridge, Massachusetts, 1985).
10. Sukarnjanaset, W., Sudhikoses, P., Tantipubthong, N., Woodley, W. L. & Rosenfeld, D. Cloud water, rainwater, temperature and draft relationships in Thai supercooled convective clouds. Preprint, *Conf. on Cloud Physics/14th Conf. on Planned and Inadvertent Weather Modification* [24–27 (American Meteorological Society, Boston, Massachusetts, 1998).
11. Williams, E. R., Zhang, R. & Rydock, J. Mixed-phase microphysics and cloud electrification. *J. Atmos. Sci.* **48**, 2195–2203 (1991).
12. Petersen, W. A. & Rutledge, S. A. On the relationship between cloud-to-ground lightning and convective rainfall. *J. Geophys. Res.* **D 103**, 14025–14040 (1996).

Acknowledgements

We thank P. Sweeney, of Weather Modification, Inc., for providing the WMI Lear jet at no charge; we also thank R. Tilbury and S. Gordon for piloting the aircraft through the vigorous Texas clouds, and K. Bosch for technical support.

Correspondence and requests for materials should be addressed to D.R. (e-mail: daniel@vms.huji.ac.il).

Past temperature and $\delta^{18}\text{O}$ of surface ocean waters inferred from foraminiferal Mg/Ca ratios

H. Elderfield* & G. Ganssen*

* Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK

† Institute of Earth Sciences, Free University, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

Determining the past record of temperature and salinity of ocean surface waters is essential for understanding past changes in climate, such as those which occur across glacial–interglacial transitions. As a useful proxy, the oxygen isotope composition ($\delta^{18}\text{O}$) of calcite from planktonic foraminifera has been shown to reflect both surface temperature and seawater $\delta^{18}\text{O}$, itself an indicator of global ice volume and salinity^{1,2}. In addition, magnesium/calcium (Mg/Ca) ratios in foraminiferal calcite show a temperature dependence^{3–5} due to the partitioning of Mg during calcification. Here we demonstrate, in a field-based calibration experiment, that the variation of Mg/Ca ratios with temperature is similar for eight species of planktonic foraminifera (when accounting for Mg dissolution effects). Using a multi-species record from the Last Glacial Maximum in the North Atlantic Ocean we found that past temperatures reconstructed from Mg/Ca ratios followed the two other palaeotemperature proxies: faunal abundance^{6,7} and alkenone saturation⁸. Moreover, combining Mg/Ca and $\delta^{18}\text{O}$ data from the same faunal assemblage, we show that reconstructed surface water $\delta^{18}\text{O}$ from all foraminiferal

species record the same glacial–interglacial change—representing changing hydrography and global ice volume. This reinforces the potential of this combined technique in probing past ocean–climate interactions.

Mg/Ca ratios have recently been used to construct a low-resolution, deep water temperature record for the Cenozoic era⁹. But there has been no validation of the method from detailed field-based calibrations for multiple species and comparison with other temperature proxies. High-resolution records obtained from planktonic foraminifera are needed to examine the response of the ocean surface to glacial–interglacial change and to provide surface ocean temperature estimates for boundary conditions of global models. Early studies of planktonic foraminifera gave conflicting evidence of a temperature effect on Mg partitioning (see, for example, ref. 10), but more recent work has yielded correlations between Mg concentrations of *Neogloboquadrina pachyderma* and annual sea surface temperature (SST) in modern core tops³ and temperature-dependent Mg uptake by *Golobigerinoides sacculifer*, *Globigerina bulloides* and *Orbulina universa* in culture^{4,5} (there is one single benthic Mg calibration based on a water depth transect¹¹), together with some initial applications of the method^{12,13}. Calibrations under culture are very valuable, but a prerequisite for palaeoceanographic applications is a demonstration that the temperature sensitivity of foraminiferal material is retained from their chamber formation through the development of gametogenic calcite—which thickens the shell at about the time of gamete release—and crust, the effects of dissolution, to residence at the sea bed. Core-top sediment

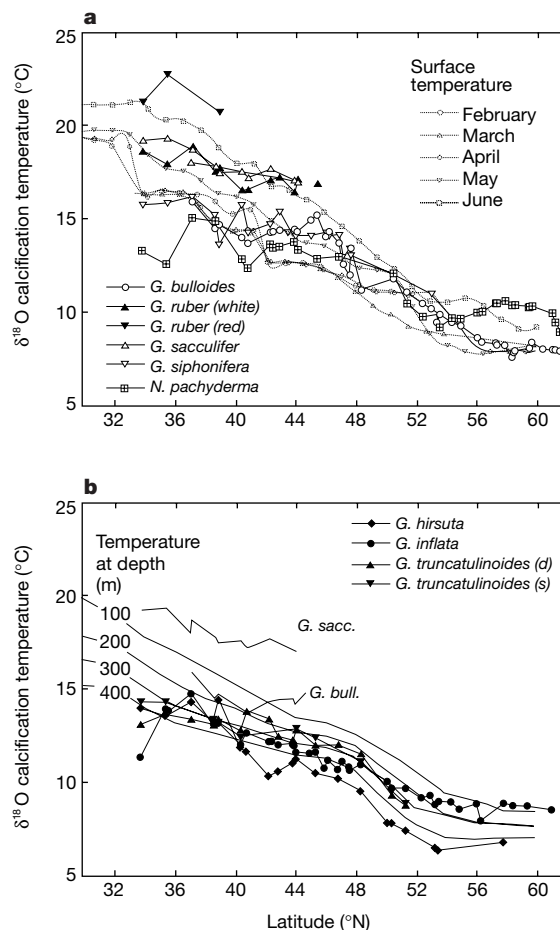


Figure 1 Foraminiferal $\delta^{18}\text{O}$ data versus latitude of core top compared with temperature data from an atlas¹⁵. **a**, Monthly surface temperatures. **b**, Annual surface temperatures at different water depths.

integrates all these processes, and is the material which passes through the sediment mixed layer to generate the palaeoceanographic record.

We analysed eight species of planktonic foraminifera on a latitudinal transect of the North Atlantic Ocean from $\sim 30^\circ$ to 60° N at about 25° W, corresponding to a range of annual SST of about 8 to 22° C. The core-top samples were radiocarbon-dated to establish their reliability¹⁴. About 20 individuals were picked for each Mg/Ca ratio analysis, and samples were analysed using a Varian Vista inductively coupled plasma optical emission spectrometer (relative precision $<0.4\%$) after sonication, cleaning, and dilute acid polishing. Details of the method are available from H.E.

Because a significant potential attraction of the method using Mg/Ca ratios is the ability to estimate temperature and seawater $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_w$) from the same foraminiferal sample, we performed paired analyses of carbonate $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_c$) and Mg/Ca ratio for our calibrations. The Mg/Ca ratio increases in the core-top samples as $\delta^{18}\text{O}_c$ decreases, consistent with a large temperature response: $\Delta(\text{Mg/Ca})/\Delta\delta^{18}\text{O}_c \approx -1 \text{ mmol mol}^{-1}$ per ‰, about a 10% change in Mg/Ca ratio per $^\circ\text{C}$. First, we have used the $\delta^{18}\text{O}_c$ data together with modern local $\delta^{18}\text{O}_w$ (obtained from a survey of $\delta^{18}\text{O}_w$

and salinity over the transect¹⁴) in the Shackleton palaeotemperature equation¹ (which we have specifically chosen because our aim is to obtain empirical Mg/Ca ratio calibrations that can be used to estimate palaeotemperatures) to calculate the mean temperatures of waters in which the foraminiferal shells were formed. When these are compared with modern ocean water temperatures¹⁵, they show the effects of both depth of calcification and seasonality (Fig. 1). For example, *N. pachyderma* and the globigerinids in general follow SST, whereas the globorotaliids (*Globorotalia hirsuta*, *Globorotalia inflata* and *Globorotalia truncatulinoides*) follow the temperature at water depths of 300–400 m. Within the globigerinids, there are seasonal effects. For example, *G. bulloides* records February to March SST between 30° and 40° N but records April to June SST at higher latitudes (see also ref. 14).

As $\delta^{18}\text{O}_c - \delta^{18}\text{O}_w$ records calcification temperature, it is logical to link Mg/Ca ratios with calcification temperature rather than with SST in an empirical calibration. When data for all species are compared, we see that all but two species to a first approximation fall on the same exponential correlation curve (Fig. 2a). The lower limits of data for the other species fall on a similar exponential correlation (Fig. 2b) but with higher values for in particular *G. bulloides* and to a lesser extent *Globigerinella siphonifera*. The use of

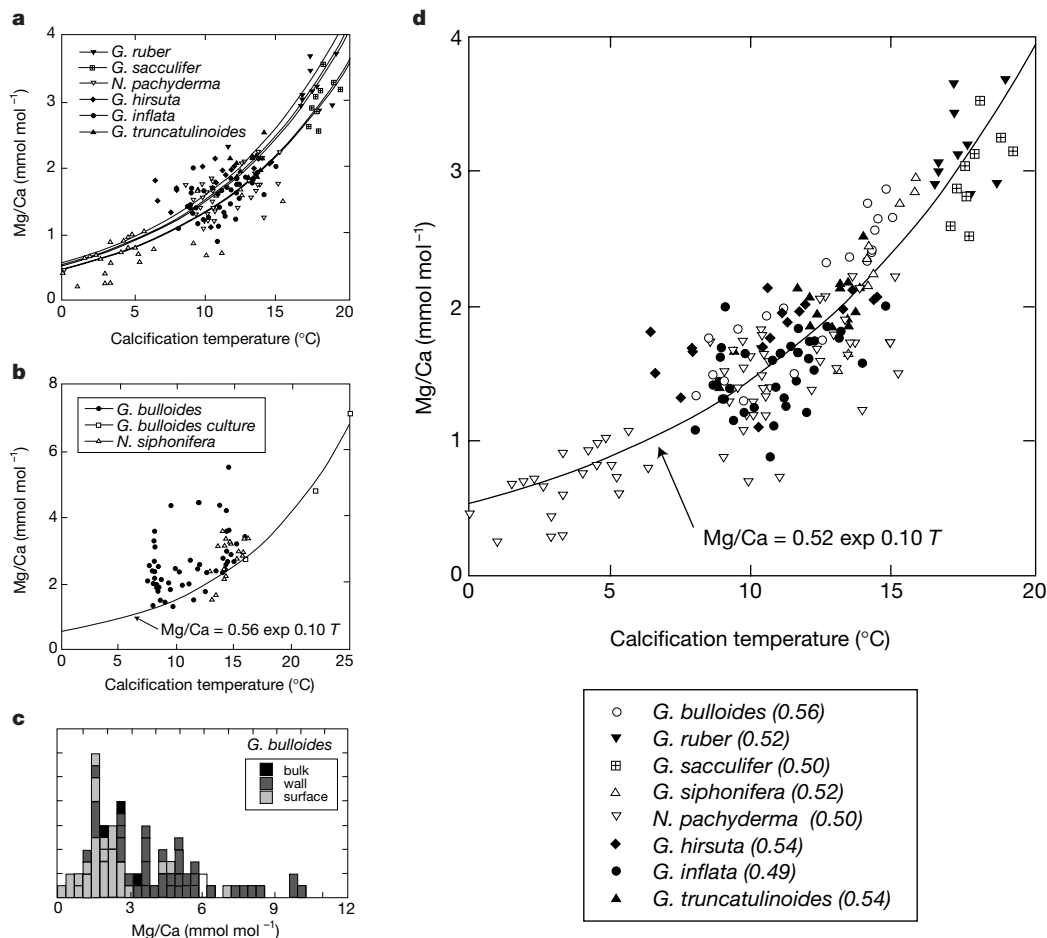


Figure 2 Mg/Ca ratios of planktonic foraminiferal calcite. **a** and **b**, Mg/Ca ratio versus calcification temperature for eight species of planktonic foraminifera from the core-top traverse. Also shown are best-fit curves for individual species and data for *G. bulloides* culture calibration⁵ curve-fitted as below. Note that *G. ruber* data are for white varieties. **c**, Electron probe microanalysis of three individual shells of *G. bulloides* from different water depths from the core-top traverse. See ref. 17 for experimental details. **d**, Mg/Ca versus calcification temperature calibration curve for eight species after correction for high-Mg carbonate phases in two species (see text and Fig. 3). Also shown are

N. pachyderma (s.) data from ref. 3. The curve shows the best fit for all data from this work and the numbers in the figure legend are the pre-exponential constants (*m*) derived from best-fit curves for individual species (**a**) of the form: $\text{Mg/Ca (mmol mol}^{-1}\text{)} = m \exp 0.10 T$, where *T* is temperature in $^\circ\text{C}$. The mean for all species gives: $\text{Mg/Ca} = 0.52 (\pm 0.0085) \exp 0.10 T$. The interspecies differences in *m* translate to temperature differences of about $\pm 0.7^\circ\text{C}$ and the standard error in the single-species core-top calibrations translates to about $\pm 0.6^\circ\text{C}$. These are similar to the value of $\pm 1.1^\circ\text{C}$ in *G. bulloides* culture data⁵.

calcification temperatures rather than estimated SST in the calibrations removes uncertainties about the depth and time of shell calcification, and provides a distinct advantage compared with calibrations based on assumed climatological SSTs.

The effect of dissolution on Mg/Ca ratios of certain planktonic foraminiferal species has already been documented^{16–18}. Therefore, we examined individuals of *G. bulloides* using an electron microprobe which show that Mg is very heterogeneously distributed within the shells (Fig. 2c). It is clear that Mg-enriched samples of *G. bulloides* and *G. siphonifera* are associated with shallower water depths (Fig. 3), and that a high-Mg phase has been removed from deeper samples, presumably by dissolution. Deciphering the origin of this high-Mg material requires further work. It cannot simply be calcite formed at a warmer temperatures because the Mg-enriched samples do not fit the calibration curves. Note also that this high-Mg material is not represented in *G. bulloides* culture (Fig. 2b). It is absent from six out of the eight species examined, and the residual calcite in all the species has about the same temperature sensitivity.

We have exploited the method of minimizing residuals in Mg/Ca ratio to refine the calibrations for individual species (Fig. 2d). The results of this approach show that, excluding samples with high residual Mg, the calibrations for each species are similar: the pre-exponential constants range between 0.49 and 0.56, mean $\pm$ s.e. is 0.52 ± 0.0085 , and the average calibration of Mg/Ca ratio (mmol mol^{-1}) versus calcification temperature ($^{\circ}\text{C}$) is $\text{Mg/Ca} = 0.52\exp(0.10T)$. This equation is similar to some literature data^{3,5}, but has a greater temperature sensitivity at higher temperatures than *G. sacculifer* culture data which were obtained by selecting low-Mg zones for analysis by electron microprobe⁴. Note that the similarities in Mg/Ca ratio between planktonic foraminifera contrast with benthic foraminifera where large interspecies differences exist which cannot be reconciled with the one existing calibration^{9,11}.

We have applied the calibrations to data for five planktonic foraminiferal species from Biogeochemical Ocean Flux Study (BOFS) core 31K recovered from the North Atlantic Ocean at $19^{\circ}00' \text{N}$, $20^{\circ}10' \text{W}$, water depth 3,200 m (see ref. 19 for details). Faunal estimates show a range between modern cold and warm SST of about 17 to 24°C , whereas modern alkenone temperatures are intermediate at about 20°C (Fig. 4a). During the last glacial period, alkenone temperatures agree with faunal T_{warm} (summer SST) but

diverge at about 9 kyr ago and this offset persists through the Holocene.

The down-core trends in $\delta^{18}\text{O}$ of the foraminifera are broadly similar but offset (Fig. 4b), as expected from Fig. 1. Mg/Ca data, translated to temperature, show a clearer separation between, on the one hand, *G. ruber* and *G. bulloides* and, on the other hand, *G. inflata*, *N. pachyderma* and *G. menardii* (Fig. 4c). Holocene temperatures for all species appear much more stable than glacial temperatures, a markedly greater glacial–interglacial contrast than seen in the faunal and alkenone estimates.

The temporal changes in temperature from Mg/Ca ratios follow roughly parallel trends for the two groups of species except for more

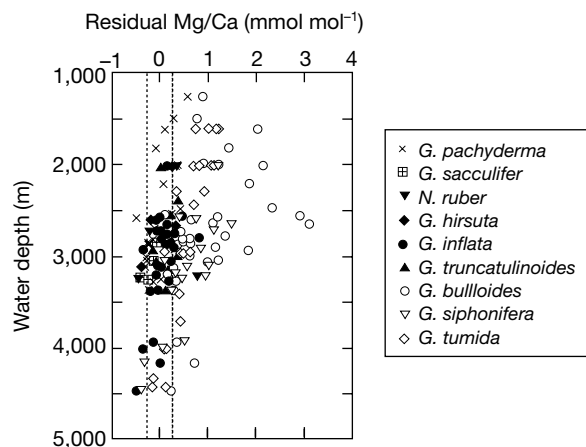


Figure 3 Residual Mg/Ca ratio versus water depth for eight species from the core-top traverse and *G. tumida* from Ontong Java Plateau depth transect¹⁷. Data were fitted to curves of the form: $\text{Mg/Ca} = m\exp(0.10T)$, and the constant m was adjusted to minimize the residual Mg/Ca ratio (measured value minus value from equation) in the deepest samples.

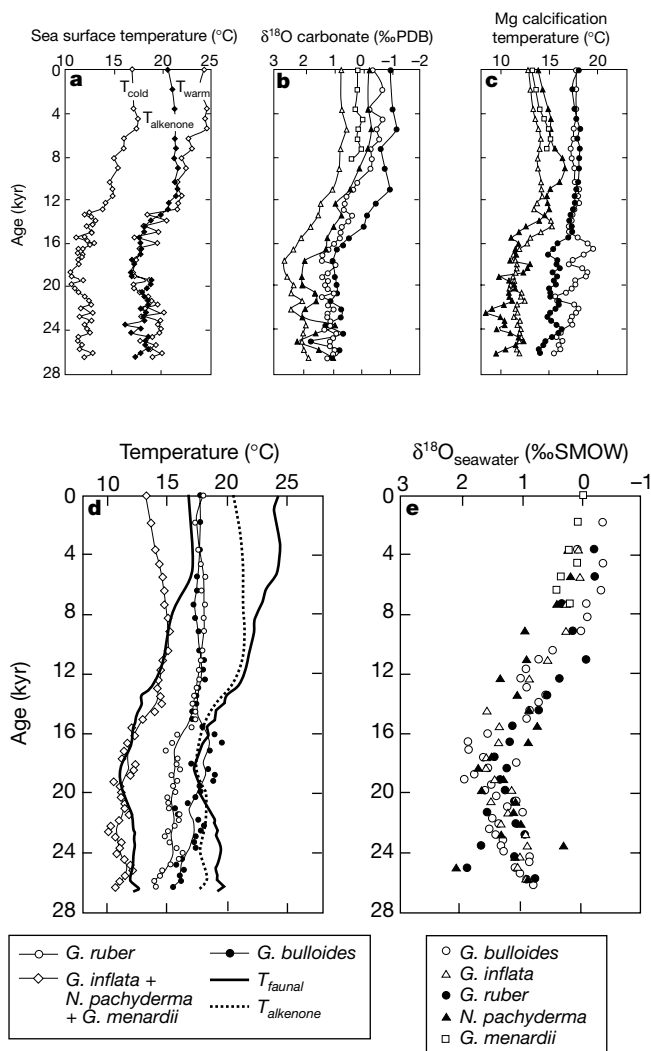


Figure 4 Records of foraminiferal oxygen isotopic compositions and Mg/Ca ratios, and temperature estimates, for BOFS core 31K. **a**, Estimates of sea surface temperature by faunal (T_{cold} and T_{warm} —winter and summer) and alkenone methods. **b**, $\delta^{18}\text{O}$ of five species of planktonic foraminifera. **c**, Calcification temperatures derived from Mg/Ca ratios using the calibration equation: $\text{Mg/Ca} (\text{mmol mol}^{-1}) = 0.52\exp(0.10T)$, where T is in $^{\circ}\text{C}$. **d**, Mg/Ca palaeotemperatures compared with faunal and alkenone estimates. For clarity, Mg/Ca data are presented as means of three foraminiferal groups: group 1, *G. ruber* (white); group 2, *G. bulloides*; group 3, *G. inflata*, *N. pachyderma* and *G. menardii*. Legend refers to **d** only. **e**, Calculated seawater $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_w$) records using foraminiferal $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_c$): $\delta^{18}\text{O}_w = \delta^{18}\text{O}_c + 0.27 - (4.38 - (4.38^2 - 0.4(16.9 - T)^{0.5}))/0.2$. Mean $\delta^{18}\text{O}_w$ increases from about $+0.8\text{‰}$ at 26 kyr ago, to 1.2 at 21 kyr ago to a maximum of $+1.5\text{‰}$ at 18 kyr ago and decreases to $+0.3\text{‰}$ at 9 kyr ago and about 0‰ at 6 kyr ago. Ages are in calendar years (see ref. 19 for details). Legend also refers to **b** and **c**.

'seasonality' or depth contrast at glacial times. The glacial temperature estimates from the three methods show similarities (in Fig. 4d *G. inflata*, *N. pachyderma* and *G. menardii* are matching cold faunal SST, and *G. ruber* and *G. bulloides* are about 2°C cooler than alkenone SST). But temperatures diverge in the Holocene. The Mg/Ca ratio tracks alkenone SST temperatures but are offset. Such differences do not necessarily mean that one method is correct or incorrect but that they define different parts of the vertical and seasonal complexity in the temperature structure of surface waters. For example, Mg/Ca temperatures for *G. ruber* are about 2.5°C cooler than alkenone SST. This may reflect intensified regional upwelling at this site which also switched the seasonal timing of maximum coccolith production away from the peak summer¹⁹.

Further work along these lines will be needed to understand these differences and thus improve all of the temperature proxies. What we can do is test for consistency and reliability of the temperatures derived by the Mg/Ca method. This can be achieved by comparing $\delta^{18}\text{O}_w$ records derived from each of the five foraminiferal species. When we do this, we see a remarkably coherent picture from all the species (Fig. 4e). The record of $\delta^{18}\text{O}_w$ at this site shows an increase from 28 kyr ago through the 6,000-calendar-year Last Glacial Maximum (LGM) interval²⁰ of 18–24 kyr ago, to a value about 1.2‰ more positive than today at the LGM. An extreme in $\delta^{18}\text{O}_w$ of about 1.5‰ more positive than today occurs at 18 kyr ago, after which $\delta^{18}\text{O}_w$ decreases through to about 9 kyr ago, after which it remains approximately constant. This pattern reflects local changes in hydrography as well as global glacial–interglacial ice volume change which is currently thought to be 1.0‰ (ref. 21). These three periods correspond well with hydrographic changes and faunal abundance variations associated with changes in intensity of trade winds and an enhanced Canary current, and migration of the Intertropical Convergence Zone¹⁹. There are differences in detail of the $\delta^{18}\text{O}_w$ records from the five species, but the broad similarity between them demonstrates the success of this approach of combining foraminiferal Mg/Ca ratios with $\delta^{18}\text{O}$ values to reconstruct both temperature and $\delta^{18}\text{O}$ of surface ocean waters. □

Received 17 September 1999; accepted 6 April 2000.

- Shackleton, N. J. Attainment of isotopic equilibrium between ocean water and the benthonic foraminiferal genus *Uvigerina*: isotopic changes in the ocean during the last glacial. *Cent. Nat. Resch. Sci. Colloq. Int.* **219**, 203–209 (1974).
- Rostek, F. et al. Reconstructing sea surface temperature and salinity using $\delta^{18}\text{O}$ and alkenone records. *Nature* **364**, 319–321 (1993).
- Nürnberg, D. Magnesium in tests of *Neoglobobulimina pachyderma* sinistral from high northern and southern latitudes. *J. Foram. Res.* **25**, 350–368 (1995).
- Nürnberg, D., Bijma, J. & Hemleben, C. Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperature. *Geochim. Cosmochim. Acta* **60**, 803–814 (1996).
- Lea, D. W., Mashiotta, T. A. & Spero, H. J. Controls of magnesium and strontium uptake in planktonic foraminifera determined by live culturing. *Geochim. Cosmochim. Acta* **63**, 2369–2379 (1999).
- CLIMAP. Seasonal reconstruction of the Earth's surface at the last glacial maximum. *Geol. Soc. Am. Map Chart Ser. MC-36*, 1–18 (1981).
- Mix, A. C., Morey, A. E., Pisias, N. G. & Hostetler, S. W. Foraminiferal faunal estimates of paleotemperature: circumventing the no-analog problem yields cool ice age tropics. *Paleoceanography* **14**, 350–359 (1999).
- Müller, P. J., Kirst, G., Ruhland, G., von Storch, I. & Rosell-Melé, A. Calibration of the alkenone paleotemperature index U_{37} based on core-tops from the eastern South Atlantic and the global ocean (60°N–60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772 (1998).
- Lear, C. J., Elderfield, H. & Wilson, P. A. Cenozoic deep sea temperatures and global ice volumes. *Science* **287**, 269–272 (2000).
- Delaney, M. L., Bé, A. W. H. & Boyle, E. A. Li, Sr, Mg and Na in foraminiferal calcite shells from laboratory culture, sediment traps, and sediment cores. *Geochim. Cosmochim. Acta* **49**, 1327–1341 (1985).
- Rosenthal, Y., Boyle, E. A. & Slowey, N. Temperature control on the incorporation of magnesium, strontium, fluorine and cadmium into benthic foraminiferal shells from Little Bahama Bank: prospects for thermocline paleoceanography. *Geochim. Cosmochim. Acta* **61**, 3633–3643 (1997).
- Hastings, D. W., Russell, A. D. & Emerson, S. R. Foraminiferal magnesium in *Globigerinoides sacculifer* as a paleotemperature proxy. *Paleoceanography* **13**, 161–169 (1998).
- Mashiotta, T. A., Lea, D. W. & Spero, H. J. Glacial-interglacial changes in Subantarctic sea surface temperature and $\delta^{18}\text{O}$ using foraminiferal Mg. *Earth Planet. Sci. Lett.* **170**, 417–432 (1999).
- Ganssen, G. M. & Kroon, D. *The Isotopic Signature of Planktonic Foraminifera from Northeast Atlantic Surface Sediments: Implications for the Reconstruction of Past Oceanic Conditions* (Special Publication, Geological Society, London, in the press).
- Levitus, S. & Boyer, T. *World Ocean Atlas 1994 Vol. 4, Temperature* (NOAA Atlas, National Environmental Satellite Data and Information Service Vol. 4, US Department of Commerce, Washington DC, 1994).

- Rosenthal, Y. & Boyle, E. A. Factors controlling the fluoride content of planktonic foraminifera: an evaluation of its paleoceanographic utility. *Geochim. Cosmochim. Acta* **57**, 355–346 (1993).
- Brown, S. J. & Elderfield, H. Variations in Mg/Ca and Sr/Ca ratios of planktonic foraminifera caused by post-depositional dissolution: evidence of shallow Mg-dependent dissolution. *Paleoceanography* **11**, 543–551 (1996).
- Rosenthal, Y., Lohmann, G. P., Lohmann, K. C. & Sherrell, R. M. Incorporation and preservation of Mg in *Gs. sacculifer*: implications for reconstructing the temperature and $^{18}\text{O}/^{16}\text{O}$ of seawater. *Paleoceanography* **15**, 135–145 (2000).
- Chapman, M. R., Shackleton, N. J., Zhao, M. & Eglinton, G. Faunal and alkenone reconstructions of subtropical North Atlantic surface hydrography and paleotemperature over the last 28 kyr. *Paleoceanography* **11**, 343–357 (1996).
- Bard, E. Ice age temperatures and geochemistry. *Science* **284**, 1133–1134 (1999).
- Schrag, D. P., Hampt, G. & Murray, D. W. Pore fluid constraints on the temperature and oxygen isotopic composition of the glacial ocean. *Science* **272**, 637–642 (1996).

Acknowledgements

We thank M. Greaves and M. Cooper for help with analyses, and D. Nürnberg, D. Lea, Y. Rosenthal, E. Bard, D. Schrag, E. Boyle and N. Shackleton for discussions.

Correspondence and requests for materials should be addressed to H.E. (e-mail: he101@esc.cam.ac.uk).

Slow differential rotation of the Earth's inner core indicated by temporal changes in scattering

John E. Vidale^{*,†}, Doug A. Dodge[‡] & Paul S. Earle^{*}

^{*} Department of Earth and Space Sciences, and [†] Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90095, USA

[‡] Lawrence Livermore National Laboratory, Geophysics Division, Livermore, California 94550, USA

The finding that the Earth's inner core might be rotating faster than the mantle¹ has important implications for our understanding of core processes, including the generation of the Earth's magnetic field^{2,3}. But the reported signal is subtle—a change of about 0.01 s per year in the separation of two seismic waves with differing paths through the core. Subsequent studies of such data have generally supported the conclusion that differential rotation exists^{4–6}, but the difficulty of accurately locating historic earthquakes⁷ and possible biases induced by strong lateral variations in structure near the core–mantle boundary⁸ have raised doubt regarding the proposed inner-core motion⁹. Also, a study of free oscillations¹⁰ constrained the motion to be relatively small compared to previous estimates and it has been proposed that the interaction of inner-core boundary topography and mantle heterogeneity might lock the inner core to the mantle¹¹. The recent detection of seismic waves scattered in the inner core¹² suggests a simple test of inner-core motion. Here we compare scattered waves recorded in Montana, USA, from two closely located nuclear tests at Novaya Zemlya, USSR, in 1971 and 1974. The data show small but coherent changes in scattering which point toward an inner-core differential rotation rate of 0.15° per year—consistent with constraints imposed by the free-oscillation data¹⁰.

We used LASA (Large Aperture Seismic Array) for this study. Although it was in operation almost three decades ago, it provides an unparalleled number and density of short-period seismometers, all buried 70 m underground^{13,14}. We examined LASA recordings of two large nuclear explosions three years apart (Table 1). The blasts were similar in size and were separated by less than 1 km (ref. 15). These were two of the four nuclear tests in the data set of 16 events used to identify inner-core scattering¹².

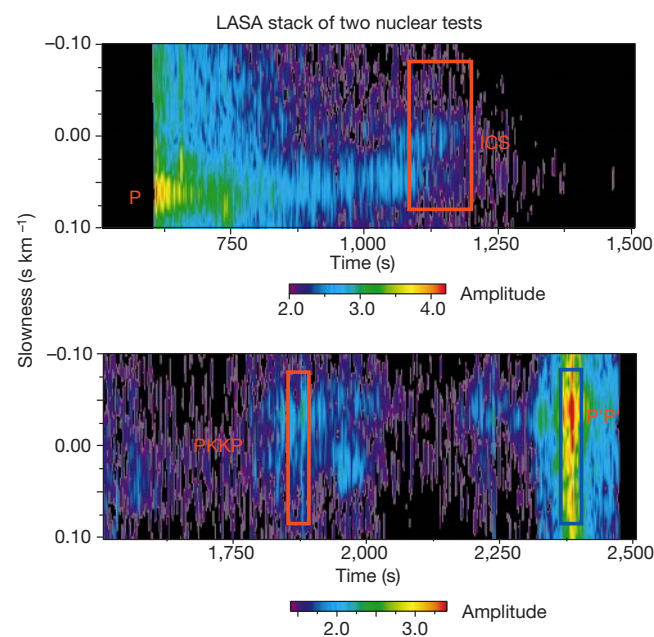


Figure 1 Slant stack of the waveform incident on LASA from the two nuclear tests. See Table 1. A slant stack displays the time and apparent slowness of incident seismic energy. The base-10 logarithm of the slant-stack amplitude envelope is shown. The upper frame shows from 608 to 1,607 s after the explosions, and the lower frame shows from 1,607 to 2,604 s. The amplitude scale is different to show weaker arrivals in the lower frame. The boxes indicate the time windows used in averaging the ICS (inner-core scattering), PKKP, and P'P' in Fig. 3 below. The strong arrival at 600 s is the direct P wave, at 650 s is the PcP wave, and at 750 s is the PP wave.

In order to build a consistent set of seismograms, we discarded any stations that did not record both explosions. This resulted in the elimination of the E and F outer rings of the LASA array, which were turned off before the 1974 blast.

Inner-core scattering (ICS) is clearest near 1 Hz (ref. 12); other phases are stronger at longer periods and the seismograms contain little teleseismic signal above 2–3 Hz. Therefore, we bandpass-filtered the data to keep a motion of only 1–2 Hz. In order to eliminate noisy traces, we cross-correlated the initial 3 s of the P wave for each station between the two events. We discarded the 21 stations with correlation coefficients less than 0.9, leaving 184 stations.

The traces are aligned with the shifts estimated from the core phases from many events¹⁶. These small shifts align the later phases better by correcting for small differences in crust and mantle structure. In general, the arrivals from the 1974 explosion are about 0.05 to 0.1 s earlier than the arrivals from 1971, probably owing to differences in near-source structure. This time difference holds all the way back to the P'P' arrival 30 min later, indicating good clock stability.

The entire waveform from 608 s to 2,604 s after the explosions is transformed to show arrival time and apparent slowness of arrivals across the array in Fig. 1. The P wave is the largest and most protracted arrival. It has a long coda that is visible until at least 1,100 s. The P arrival is clipped: the true ground motion for the direct P wave was much larger than shown. The ICS wave is clearly visible from 1,100 s until 1,250 s near zero slowness. PKKP, a wave that reflects from the underside of the core–mantle boundary, appears as a complicated series of arrivals with a range of slownesses from 1,800 to 2,000 s. A strong P'P' arrival is seen near 2,350 s. This phase reflects from the surface at the far side of the Earth.

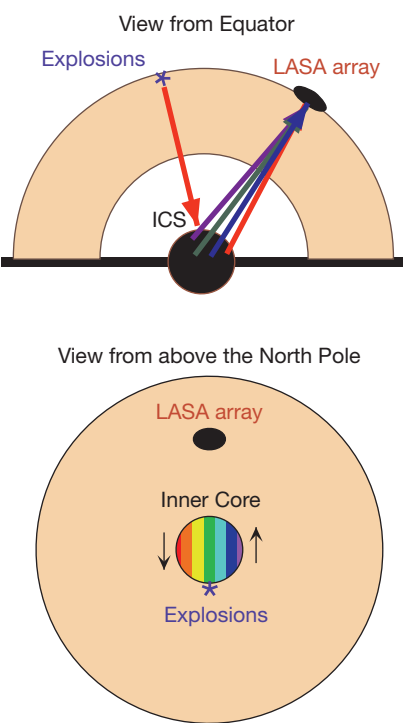


Figure 2 Cartoon of inner-core scattering path. The upper cartoon shows the view from the plane of the Equator. We model ICS arrivals as scattering from fine-scale heterogeneity uniformly distributed in the inner core. The lower cartoon shows the view from above the North Pole. Counterclockwise (faster than the mantle) inner-core differential rotation would lead to the blue part of the inner core advancing toward LASA and the red part of the inner core receding from LASA over time.

Our analysis hinges on interpreting the variation in the ICS waves in terms of relative inner-core rotation. A cartoon of the paths of these scattered waves is shown in Fig. 2. Our geometry may be approximated by considering explosions on the North Pole and the array in its true location in Montana. We only see the outermost quarter of the inner core with ICS because of its strong attenuation. If the inner core rotates faster than the mantle, then the structures in the inner core to the west of LASA (blue in Fig. 2) will come nearer over time. Structures in the region to the east (red in Fig. 2) will recede. Thus, inner-core rotation would lead to earlier scattered arrivals from the west in 1974 compared to 1971, and later scattered arrivals from the east.

This pattern of advancing arrivals from the west and delayed arrivals from the east is seen in Fig. 3. Between 1971 and 1974, there is a change of more than 0.1 s in the arrival times of the ICS as a function of transverse slowness. An alternative we must consider is the possibility that the spatial separation of the two explosions causes the observed variation in the ICS waves. A possible example of the effect of source separation is the red colour of the late P coda energy that appears at greater radial slowness with the ICS. However, the P coda leaves the source region at shallower angles than the ICS waves, and thus may have a greater time shift from the source separation. We therefore next compare all three phases that have traversed the core: ICS, PKKP and P'P'. The radial and transverse

Table 1 Nuclear explosions					
Date	Time	Latitude (°N)	Longitude (°E)	Depth (km)	Distance
27 Sept. 1971	05:59:55.75	73.393	54.929	0	59.5°
29 Aug. 1974	09:59:56.15	73.397	54.914	0	59.5°

These locations are from ref. 15.

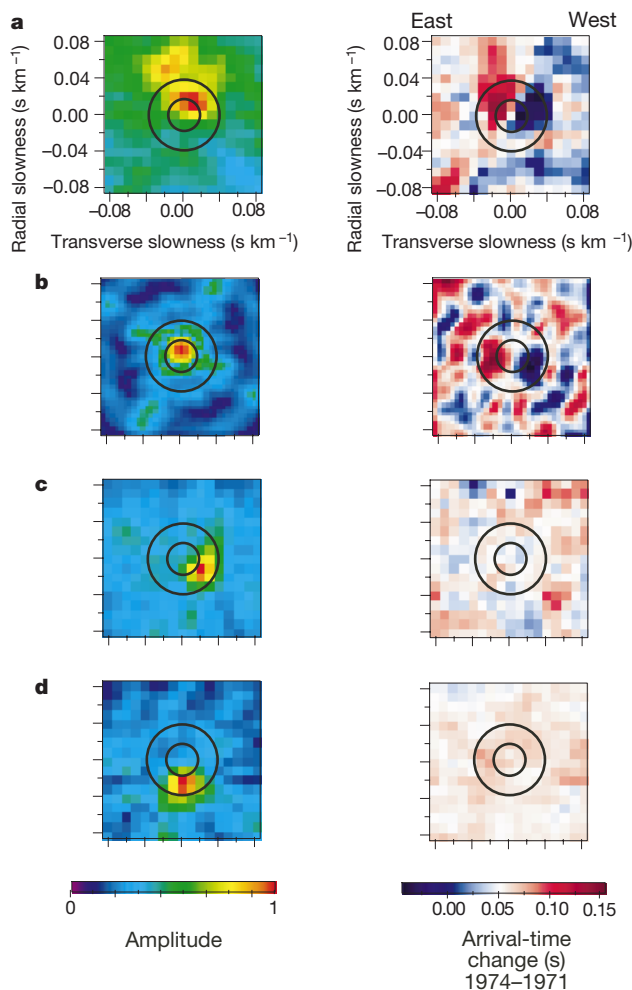


Figure 3 Amplitude and time difference as a function of radial and transverse slowness. Seismic waves that are vertically incident, for example, appear in the middle of such a plot, at 0.0 km s^{-1} transverse and radial slowness. The ring of rays that would graze the inner-core boundary and the core–mantle boundary are shown by the black circles in each frame. **a–d**, Left side shows the distribution of amplitude for ICS (**a**), simulated ICS with 0.45° rotation calculation (**b**), PKKP (**c**) and P'P' (**d**). We convolved the beam response in the simulation to mimic how well LASA resolves the slowness of the incident waves. ICS is averaged from 1,100 to 1,200 s, PKKP from 1,860 to 1,890 s, and P'P' from 2,360 to 2,390 s. Right side shows the distribution of change in arrival time between 1971 and 1974 as a function of radial and transverse slowness for the same phases (**a–d**). The phase delays are an amplitude-weighted average across the time windows. The crucial feature is that ICS arrivals from the east (left) are later in 1974 than in 1971, and that ICS arrivals from the west (right) are earlier. The PKKP and P'P' arrivals show much less difference between 1971 and 1974. Axes are as labelled in **a**.

slowness of the arrivals, visible in Fig. 3, verifies that all three arrivals impinge on LASA from steep angles.

The ICS phase shows more variation in delay with azimuth than the PKKP and P'P' arrivals in Fig. 3. This is particularly true for the regions of slowness space in which there is considerable amplitude. P'P' is well known to exhibit a strong component of scattering¹⁷, and PKKP displays arrivals from a wide range of back-azimuths, indicating very strong scattering¹⁸. Thus, all three phases are expected to contain energy from a range of azimuths. In addition, all three phases may contain energy that samples the inner core. However, ICS waves are reflected from within the inner core, while PKKP and P'P' waves mostly miss the inner core and those waves that penetrate the inner core mostly have transmitted paths. Thus, changes in PKKP and P'P' from inner-core rotation should be

negligible, as we observed. It is most likely that the small separation between the two explosions contributes very little to the differences in ICS shown in Fig. 3.

Our calculation¹² assumes Born single scattering and uniformly distributed 1.2% variations in the moduli and density with a 2-km wavelength, attenuation Q_p of 360 (ref. 19), and traces rays through PREM (ref. 20). The results are not sensitive to these choices. We assume differential rotation about the same pole as the rotation of the Earth. Movement of the inner core would change the scattered wavefield in two ways: focusing and interference patterns would change, and the arrival times of the scattered energy would shift in time. We ignore the former effect, which would not lead to systematic time shifts, and calculate the latter, which would.

To calculate the average arrival-time shift of the scattered wave-train, we assume all points within the inner-core contribute to the observed ICS. The time shift resulting from a given amount of differential rotation for a given scatterer (point in the inner-core) is calculated by finding the difference between the travel-time to the scatterer before and after the rotation. This delay is weighted by its amplitude contribution to the entire ICS wavetrain. The amplitude contribution for a particular scatterer is affected by geometrical spreading, inner-core attenuation, scattering angle, and the contrast in elastic properties.

Figure 3 also shows the predicted amplitude and time delay patterns resulting from an inner-core rotation of 0.45° , given our scattering model. These model patterns match the observations well. The observed time delay of $\pm 0.1 \text{ s}$ accrued over three years determines that the inner core was rotating 0.15° per year faster than the mantle. Some differences are visible, as well. Late P coda energy at the greater radial slowness of 0.05 s km^{-1} appears in the amplitude map of the ICS, and was present for both explosions. This energy is also visible in Fig. 1 as the streak with the P wave slowness. In contrast, there is no P coda in the calculation. Another difference is the presence of background noise in the data but not in the calculated amplitude plot. The lack of background noise in the calculation leads to the significant time-delay oscillations in the arrival-time change plot at slownesses with little power.

We confirm that the inner-core rotates as first claimed in 1996 (ref. 1) but at a considerably slower rate: 0.15° per year compared to $0.5\text{--}1.1^\circ$ per year. Our estimate falls within the bounds of $0 \pm 0.2^\circ$ per year set by some measurements of differential travel times⁷ and normal modes¹⁰. Future studies of ICS waves can be used to search for time-dependence in the differential inner-core rotation and to test the assumption that the differential rotation is about the Earth's north–south axis. □

Received 25 October 1999; accepted 31 March 2000.

1. Song, X. D. & Richards, P. G. Seismological evidence for differential rotation of the Earth's inner core. *Nature* **382**, 221–224 (1996).
2. Glatzmaier, G. A. & Roberts, P. H. A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature* **377**, 203–209 (1995).
3. Kuang, W. L. & Bloxham, J. An earthlike numerical dynamo model. *Nature* **389**, 371–374 (1997).
4. Song, X. Joint inversion for inner core rotation, inner core anisotropy, and mantle heterogeneity. *J. Geophys. Res.* **105**, 7931–7944 (2000).
5. Creager, K. C. Inner core rotation rate from small-scale heterogeneity and time-varying travel times. *Science* **278**, 1284–1288 (1997).
6. Song, X. & Li, A. Support for differential inner core superrotation from earthquakes in Alaska recorded at South Pole station. *J. Geophys. Res.* **105**, 623–630 (2000).
7. Poupinet, G., Souriau, A. & Coutant, O. The existence of an inner core super-rotation questioned by teleseismic doublets. *Phys. Earth Planet. Inter.* **118**, 77–88 (2000).
8. Breger, L. & Romanowicz, B. Three-dimensional structure at the base of the mantle beneath the central Pacific. *Science* **282**, 718–720 (1998).
9. Souriau, A., Roudil, P. & Moynot, B. Inner core differential rotation: Facts and artefacts. *Geophys. Res. Lett.* **24**, 2103–2106 (1997).
10. Laske, G. & Masters, T. G. Rotation of the inner core from a new analysis of free oscillations. *Nature* **402**, 3397–3400 (1999).
11. Buffett, B. A. Geodynamic estimates of the viscosity of the Earth's inner core. *Nature* **388**, 571–573 (1997).
12. Vidale, J. E. & Earle, P. S. Fine-scale heterogeneity in the Earth's inner core. *Nature* **404**, 273–275 (2000).
13. Green, P. E., Frosch, R. A. & Romney, C. F. Principles of an experimental large aperture seismic array (LASA). *Proc. IEEE* **53**, 1821–1833 (1965).

14. Capon, J. Analysis of Rayleigh-wave multipath at LASA. *Bull. Seismol. Soc. Am.* **60**, 1701–1731 (1970).
15. Marshall, P. D., Porter, D., Young, J. B. & Peachall, P. A. Atomic Weapons Establishment Report O 2/94 (Her Majesty's Stationery Office, London, 1994).
16. Engdahl, E. R. & Felix, C. P. Nature of travel-time anomalies at Lasca. *J. Geophys. Res.* **76**, 2706–2715 (1971).
17. Haddon, R. A. W., Husebye, E. S. & King, D. W. Origin of precursors to P'P'. *Phys. Earth Planet. Inter.* **14**, 41–70 (1977).
18. Earle, P. S. & Shearer, P. M. Observations of high-frequency scattered energy associated with the core phase PKKP. *Geophys. Res. Lett.* **25**, 405–408 (1998).
19. Bhattacharyya, J., Shearer, P. M. & Masters, T. G. Inner core attenuation from short period PKP(BC) versus PKP(DF) waveforms. *Geophys. J. Int.* **114**, 1–11 (1993).
20. Dziewonski, A. M. & Anderson, D. L. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).

Acknowledgements

We thank the Albuquerque Seismological Laboratory and B. Woodward and H. Bolton for access to LASA data. Discussions with K. Creager, P. Richards, T. Wallace, A. Souriau and P. Davis were helpful.

Correspondence and requests for materials should be addressed to J. E. V. (e-mail: vidale@ucla.edu).

Invasion sequence affects predator–prey dynamics in a multi-species interaction

Steven M. Sait*, Wei-Chung Liu†, David J. Thompson*, H. Charles J. Godfray† & Michael Begon*

* Population and Evolutionary Biology Research Group, School of Biological Sciences, Nicholson Building, The University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

† Department of Biology and NERC Centre for Population Biology, Imperial College at Silwood Park, Ascot SL5 7PY, UK

Ecologists seek to understand the rules that govern the assembly, coexistence and persistence of communities of interacting species. There is, however, a variety of sequences in which a multi-species community can be assembled—unlike more familiar one- and two-species systems. Ecological systems can exhibit contrasting dynamics depending on initial conditions¹, but studies have been focused on simple communities initiated at different densities, not on multi-species communities constructed in different sequences. Investigations of permanence and convergence in ecological communities^{2–4} have been concerned with the flux of whole species (presence or absence)⁴ but have not addressed the central issues concerning the dynamics exhibited by individual species in particular interactions. Here we examine data for replicated three-species systems and demonstrate that the dynamic trajectories of both a predator and its prey within the system are determined by the sequence in which it is constructed, and that for one construction-sequence alternative dynamic patterns are possible.

The multi-species system investigated here comprises a lepidopteran host (the moth *Plodia interpunctella*) and two of its natural enemies: a viral pathogen and a parasitoid wasp (a type of predator). In continuous culture, either alone or with the pathogen, the host exhibits 'generation cycles' of abundance⁵, with periods of between six and seven weeks^{6,7}. Populations coexisting with the pathogen have high host densities (similar to those in host-alone cultures) and a pathogen prevalence of less than 10%. Intra-specific competition for resources drives these cyclic dynamics; the pathogen simply modulates the host's vital rates (development time and reproduction)^{8,9}. In host–parasitoid cultures, both species also display regular generation cycles, but the abundance of the host is

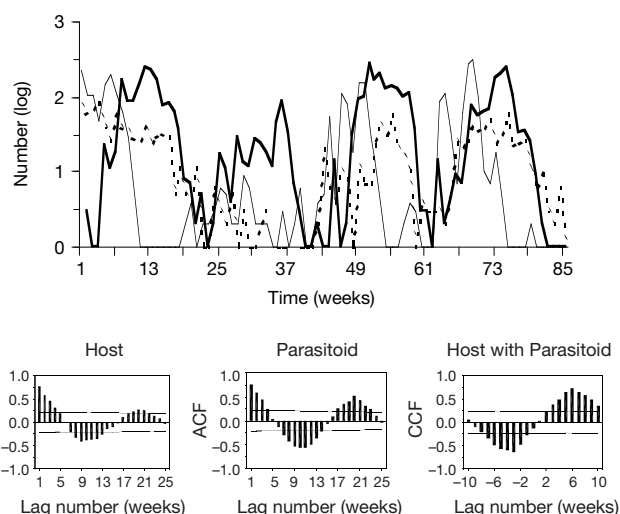


Figure 1 Representative culture of *P. interpunctella* (thin line) and granulovirus (GV) (dotted line) to which *V. canescens* (bold line) was added. The autocorrelation function analyses (ACFs, bottom) of the host and parasitoid demonstrate an immediate shift from generation to multi-generation cycles, whereas the cross-correlation function analysis (CCF) shows a significant lag (approximately 1 generation) between host and parasitoid fluctuations. The solid bars represent the correlation coefficients, the sloping lines represent confidence intervals of two standard errors, which identify significant lags and their periods in weeks ($P < 0.05$).

significantly reduced and the amplitude of the cycles greatly enhanced¹⁰; it is a truly trophic interaction in which both predator and prey exhibit a numerical response to one another's densities¹¹.

We have shown previously that the three-species system constructed by adding the parasitoid to an established host–virus system (referred to as host–virus–parasitoid, HVP) exhibits dynamics strikingly different from those of its two-species counterparts (Fig. 1): multi-generation cycles in the host and parasitoid and eventual extinction of the entire system⁷. Furthermore, cross-correlation function (CCF) analysis demonstrates that the host and parasitoid were not in phase, unlike two-species host–parasitoid interactions (see Supplementary Information). Instead, there was a significant, approximately one-generation lag, indicative of a delayed density-dependent response of the parasitoid to host abundance (Fig. 1).

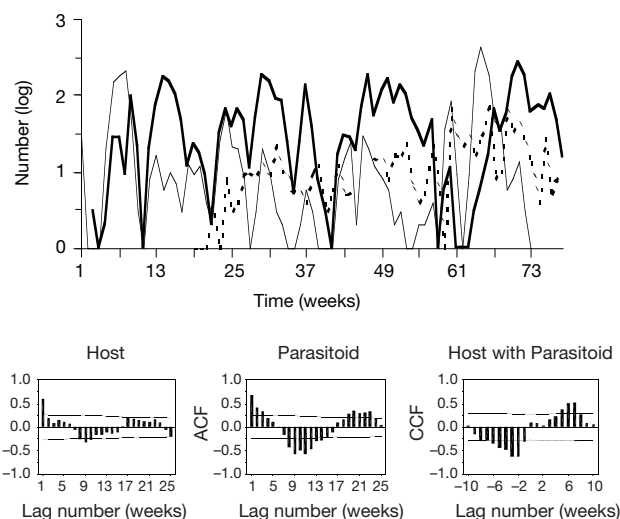


Figure 2 Representative culture of *P. interpunctella* and *V. canescens* to which GV was added. In these replicates, there was also a shift to the multi-generation cycle (ACFs) and again there was a lag between host and parasitoid dynamics (CCF) (see Supplementary Information).

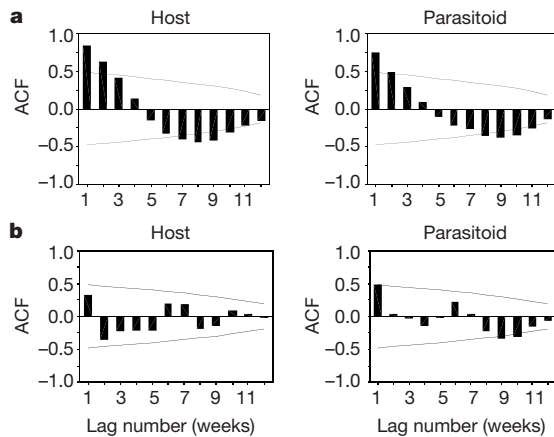


Figure 3 ACFs of the host and parasitoid for the first 14 weeks (approximately two host generations). **a**, For the HVP population, there was an immediate breakdown in the generation cycle (no evidence of generation-length peaks) and shift to multi-generation cycles, the ACFs being characteristic of an increasing trend in abundance²³ (to be expected over the initial phase of a cycle 3–4 host generations in length⁷). **b**, The HPV population exhibiting multi-generation cycles. The host and parasitoid exhibit a near-significant peak at the generation cycle but no evidence of trend, demonstrating a gradual shift to the multi-generation pattern (see Supplementary Information).

When, by contrast, the virus was added to six established host–parasitoid populations (that is, HPV systems were constructed) the host and parasitoid exhibited either one of two dynamic patterns, neither the same as that of the HVP system. In the first, displayed by three of six replicates, the host and parasitoid both exhibited multi-generation cycles with a one-generation lag between them (Fig. 2), but the transition from generation to multi-generation cycles was only gradual, contrasting markedly with the immediate shift to a multi-generation cycle in the HVP systems (Fig. 3).

The second, contrasting pattern was displayed by the other three replicates, where there was no evidence of either a shift to multi-generation dynamics or of a persisting generation cycle (Fig. 4). However, despite the lack of a regular cyclical pattern, the host and parasitoid dynamics remained in phase (Fig. 4). Thus, despite the subsequent breakdown of the host generation cycle following the

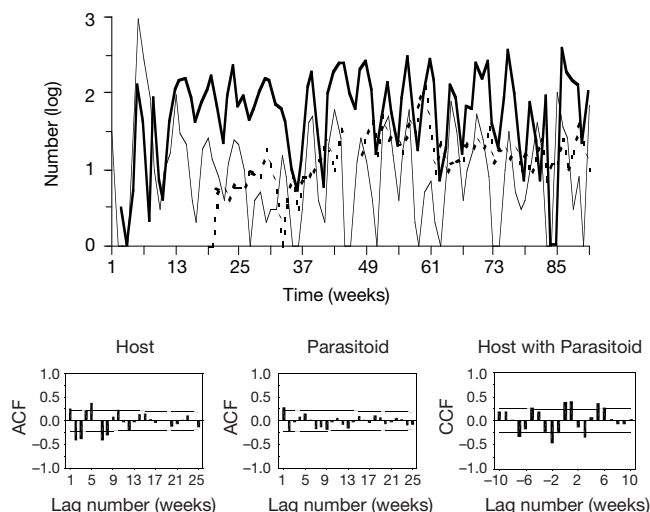


Figure 4 Representative culture of *P. interpunctella* and *V. canescens* to which GV was added in which there was no significant pattern in the host and parasitoid dynamics (ACFs), although they remained in-phase (CCF shows a significant correlation at lag 0) (see Supplementary Information).

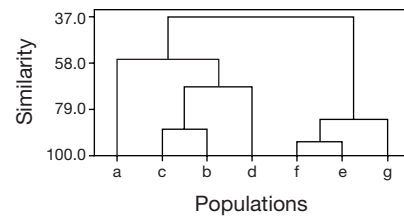


Figure 5 Similarity dendrogram following hierarchical cluster analysis of HVP and HPV cultures. The populations exhibiting multi-generation cycles cluster into one group; those without a significant pattern form another. Moreover, the group of multi-generation HPV cultures is also closer to but quite distinct from the HVP culture, reflecting the different rates at which these populations reach the multi-generation pattern. Populations a, b and e refer to Figs 1, 2 and 4 respectively; c and d refer to the remaining two replicates that exhibited multi-generation cycles; f and g refer to those replicates that did not exhibit a pattern (see Supplementary Information).

addition of the pathogen, the host and parasitoid maintained their synchronized development patterns and the parasitoid was able to respond in an immediate density-dependent manner to changes in host abundance. This clear distinction between the cultures in Figs 1 and 2, and that in Fig. 4 is further emphasized by a cluster analysis of their autocorrelation function (ACF) and CCF characteristics (Fig. 5).

Models of ecological systems show that alternative equilibrium attractor states are possible, with the outcome dependent on initial densities or densities stochastically arrived at¹. However, theoretical investigations^{12,13} have typically focused on conditions for coexistence and on equilibrium dynamics, whereas our results join others^{1,14–16} indicating that transient dynamics may be as or more important. In the communities described here, the HVP systems

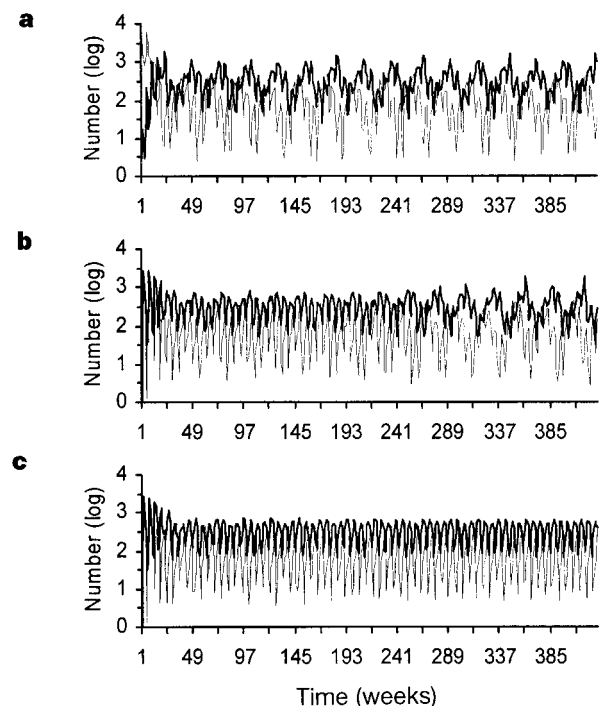


Figure 6 Patterns exhibited by the three-species age-structured model. Outputs from the model of the HVP (**a**) and HPV (**b,c**) systems. For each construction sequence the parameter values are the same, except the virus transmission coefficient is 0.08 in **a** (only one HVP trajectory is given here but all outputs exhibit the same pattern) and is 0.06 and 0.08 in **b** and **c**, respectively. Only host (thin line) and parasitoid (bold line) trajectories are shown.

and half the HPV systems moved either rapidly (HVP) or slowly (HPV) to a pattern of multi-generation predator–prey cycles. We could not determine whether, ultimately, all of the remaining HPV systems would exhibit such behaviour as two of the three replicates rapidly went extinct. However, one of the populations (Fig. 4) persisted, and failed to exhibit cycles, well beyond the period in which other HPV replicates exhibited multi-generational cycles (roughly 90 weeks compared with typically 21 weeks; Fig. 2), convincing us that ‘transient’ behaviour may be long enough here to be significant ecologically.

An age-structured model of this host–pathogen–parasitoid interaction also shows different dynamic behaviour depending on the order in which the three-species system is assembled (Fig. 6). When the parasitoid is added to the host–pathogen system, multi-generation cycles occur rapidly, an output that is robust to a wide range of parameter values (Fig. 6a). When the pathogen is added to the host–parasitoid system, however, multi-generation cycles appear again—but more slowly (Fig. 6b). Moreover, the length of the transient period before multi-generation cycles is sensitive to small changes in pathogen density in the period following invasion, which can itself be reduced simply by reducing the transmission coefficient of the pathogen (Fig. 6b and c). As with the data, the length of the transient can be so long (Fig. 6c) that it may be indistinguishable from what appears to be an alternative stable state.

What could be the biological mechanism underlying the contrasting dynamics in these experimental systems? Established host–pathogen systems are characterized by high host and pathogen density^{6,8,17}. When the parasitoid invades, the abundance of both natural enemies is high almost immediately. Susceptibility to pathogen infection is much greater in the early larval stages¹⁸ whereas the parasitoid actively prefers the later stages¹⁹. Thus, the combined attack from both exploits all larval stages. Conventional predator–prey models predict multi-generation cycles in such circumstances⁷. In host–parasitoid systems, on the other hand, high parasitoid abundance severely suppresses host numbers, and low larval density markedly reduces cannibalism, the major route of pathogen transmission²⁰. Thus, the virus can invade only slowly. Therefore, because the principle mortality factor is the parasitoid acting in a stage-dependent manner, which is predicted to promote generation cycles^{21,22}, additional limited pathogen-induced mortality may simply cause the predator–prey generation cycles to be blurred: a remnant of the generation cycle persists, the host and parasitoid cross-correlate in phase, and subsequent changes in pathogen abundance over time result in some of the HPV systems embarking on the multi-generation cycle trajectory.

Species invasions of established communities are pervasive throughout ecology: some invasions natural, others anthropogenic. Our study indicates that understanding the population dynamic consequences of the sequence in which a multi-species interaction is assembled will be essential for predicting the outcome of invasions, and that the ‘transient’ behaviour of such systems following an invasion may be as, or more, important than equilibrium states that are effectively never attained over the timescales observed in complex natural systems. □

Methods

The three species

The host has five larval instars (stages) and a generation time of around 42 days⁶. The pathogen is a granulovirus, a member of the baculovirus family, which only infects host larvae following ingestion of virus particles; the lethal dose increases with larval instar¹⁸. The parasitoid, *Venturia canescens*, lays eggs singly in moth larvae (preferentially later instars¹⁹). The developing parasitoid synchronises its growth with that of the host, consuming the final instar before emerging as an adult²⁴. The census data from the populations, which were logged before analysis, were recorded as a weekly count of dead adult moths and wasps and, where appropriate, the prevalence of infection was recorded as the number of infected larvae. HVP populations were initiated by adding two adult wasps to established host–pathogen populations⁷. HPV cultures were initiated by adding twelve virus-infected larvae to established host–parasitoid systems.

Cluster analysis

Three independent variables were calculated from the ACF and CCF outputs of the host in each of the three species trajectories; value of the ACF at the generation cycle lag (mean of lags 5, 6 and 7); value of the ACF at the multi-generation cycle lag (mean of lags 19, 20 and 21); ratio of the CCF at lag 0 to that at lag 6 (indicative of the strength of in-phase density-dependence or delayed density-dependence between host and parasitoid abundance).

The model

We model this highly age-structured three-species interaction as systems of time-delayed ordinary differential equations using the lumped age-class method²³, and incorporate parameter values measured in this system where possible; both host–pathogen and host–parasitoid models display generation cycles as observed in the laboratory. Host larvae compete for resources, with older larvae having a disproportionate effect on young larvae, and with egg cannibalism occurring (see ref. 17 for a full description and analysis of the single species host model). Larval hosts are divided into two age classes¹⁷, both of which are susceptible to the pathogen and parasitoid. Young hosts are killed by the pathogen but older hosts develop a non-lethal infection that reduces viability and fecundity². Parasitoid attack rates on older larvae are greater than on younger larvae, and parasitoid development is delayed in younger larvae^{19,24}. A nonlinear attack/transmission function is assumed, and the natural enemies incorporated as described in previous studies^{21,26}.

Received 10 January; accepted 17 March 2000.

- Cushing, J. M., Dennis, B., Desharnais, R. A. & Costantino, R. F. Moving towards an unstable equilibrium: saddle nodes in population systems. *J. Anim. Ecol.* **67**, 298–306 (1998).
- Law, R. & Morton, R. D. Permanence and the assembly of ecological communities. *Ecology* **77**, 762–775 (1996).
- Samuels, C. L. & Drake, J. A. Divergent perspectives on community convergence. *Trends Ecol. Evol.* **12**, 427–432 (1997).
- Law, R. in *Advanced Ecological Theory* (ed. McGlade, J.) 143–171 (Blackwell Science, Oxford, 1999).
- Gurney, W. S. C. & Nisbet, R. M. Fluctuation periodicity, generation separation and the expression of larval competition. *Theor. Popul. Biol.* **28**, 150–180 (1985).
- Sait, S. M., Begon, M. & Thompson, D. J. Long-term population dynamics of the Indian meal moth *Plodia interpunctella* and its granulosis virus. *J. Anim. Ecol.* **63**, 861–870 (1994).
- Begon, M., Sait, S. M. & Thompson, D. J. Predator–prey cycles with period shifts between two- and three-species systems. *Nature* **381**, 311–315 (1996).
- Bjørnstad, O. N. et al. Population dynamics of the Indian meal moth: demographic stochasticity and delayed regulatory mechanisms. *J. Anim. Ecol.* **67**, 110–126 (1998).
- Sait, S. M., Begon, M. & Thompson, D. J. The effect of a sublethal baculovirus infection in the Indian meal moth, *Plodia interpunctella*. *J. Anim. Ecol.* **63**, 541–550 (1994).
- Begon, M., Sait, S. M. & Thompson, D. J. Persistence of a parasitoid–host system: refuges and generation cycles? *Proc. R. Soc. Lond. B* **260**, 131–137 (1995).
- Bjørnstad, O. N., Sait, S. M., Stenseth, N. C., Thompson, D. J. & Begon, M. Coupling strength controls the impact of specialised enemies on the dimension complexity of experimental moth dynamics. *Nature* (submitted).
- May, R. M. & Hassell, M. P. The dynamics of multiparasitoid–host interactions. *Am. Nat.* **117**, 234–261 (1981).
- Hochberg, M. E., Hassell, M. P. & May, R. M. The dynamics of host–parasitoid–pathogen interactions. *Am. Nat.* **135**, 74–94 (1990).
- Rand, D. A. & Wilson, H. B. Chaotic stochasticity—a ubiquitous source of unpredictability in epidemics. *Proc. R. Soc. Lond. B* **246**, 179–184 (1991).
- Hastings, A. & Higgins, K. Persistence of transients in spatially structured ecological models. *Science* **263**, 1133–1136 (1994).
- Dennis, B., Desharnais, R. A., Cushing, J. M. & Costantino, R. F. Transitions in population dynamics: equilibria to periodic cycles to aperiodic cycles. *J. Anim. Ecol.* **66**, 704–729 (1997).
- Briggs, C. J., Sait, S. M., Begon, M., Thompson, D. J. & Godfray, H. C. J. What causes generation cycles in populations of stored product moths? *J. Anim. Ecol.* **69**, 352–366 (2000).
- Sait, S. M., Begon, M. & Thompson, D. J. The influence of larval age on the response of *Plodia interpunctella* to a granulosis virus. *J. Invertebr. Pathol.* **63**, 107–110 (1994).
- Sait, S. M., Begon, M., Thompson, D. J., Harvey, J. A. & Hails, R. S. Factors affecting host selection in an insect host–parasitoid interaction. *Ecol. Entomol.* **22**, 225–230 (1997).
- Knell, R. J., Begon, M. & Thompson, D. J. Transmission of *Plodia interpunctella* granulosis virus does not conform to the mass action model. *J. Anim. Ecol.* **67**, 592–599 (1998).
- Godfray, H. C. J. & Hassell, M. P. Discrete and continuous insect populations in tropical environments. *J. Anim. Ecol.* **58**, 153–174 (1989).
- Gordon, D. M., Nisbet, R. M., De Roos, A., Gurney, W. S. C. & Stewart, R. R. Discrete generations in host–parasitoid models with contrasting life cycles. *J. Anim. Ecol.* **60**, 295–308 (1991).
- Chatfield, C. *The Analysis of Time Series* (Chapman & Hall, London, 1996).
- Harvey, J. A., Harvey, I. F. & Thompson, D. J. Flexible larval growth allows use of a range of host sizes by a parasitoid wasp. *Ecology* **75**, 1420–1428 (1994).
- Gurney, W. S. C., Nisbet, R. M. & Lawton, J. H. The systematic formulation of tractable single-species population models incorporating age structure. *J. Anim. Ecol.* **52**, 479–495 (1983).
- Briggs, C. J. & Godfray, H. C. J. The dynamics of insect–pathogen interactions in stage-structured populations. *Am. Nat.* **145**, 855–887 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank the NERC for funding this work, P. Rohani and O. Bjørnstad for helpful comments and constructive criticism of earlier versions of the manuscript.

Correspondence and requests for material should be addressed to S.M.S. (e-mail: lesait@liv.ac.uk).

Chromosomal evolution in *Saccharomyces*

G. Fischer*, S. A. James†, I. N. Roberts‡, S. G. Oliver‡ & E. J. Louis*

* Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

† National Collection of Yeast Cultures, Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, UK

‡ School of Biological Sciences, University of Manchester, 2.205 Stopford Building, Oxford Road, Manchester M13 9PT, UK

The chromosomal speciation model invokes chromosomal rearrangements as the primary cause of reproductive isolation¹. In a heterozygous carrier, chromosomes bearing reciprocal translocations mis-segregate at meiosis, resulting in reduced fertility or complete sterility. Thus, chromosomal rearrangements act as a post-zygotic isolating mechanism. Reproductive isolation in yeast is due to post-zygotic barriers, as many species mate successfully but the hybrids are sterile^{2,3}. Reciprocal translocations are thought to be the main form of large-scale rearrangement since the hypothesized duplication of the whole yeast genome 10⁸ years ago^{4,5}. To test the chromosomal speciation model in yeast, we have characterized chromosomal translocations among the genomes of six closely related species in the *Saccharomyces* 'sensu stricto' complex⁶. Here we show that rearrangements have occurred between closely related species, whereas more distant ones have colinear genomes. Thus, chromosomal rearrangements are not a prerequisite for speciation in yeast and the rate of formation of translocations is not constant. These rearrangements appear to result from ectopic recombination between Ty elements or other repeated sequences.

Potential mechanisms of speciation include the accumulation of major gene differences, chromosomal rearrangements and simple sequence divergence upon which the mismatch repair system acts. The relative importance and contributions of different mechanisms leading to speciation can be measured in appropriate taxa for which genomic and genetic data are available. The *Saccharomyces* yeasts give us an unparalleled opportunity to make these measurements.

Saccharomyces 'sensu stricto' is an evolutionary lineage distinct from the species belonging to the *Saccharomyces* 'sensu lato' group^{2,7}. The sensu stricto group comprises *S. cerevisiae*, *S. paradoxus*, *S. bayanus* and *S. pastorianus*, as well as three newly defined species, *S. cariocanus*, *S. mikatae* and *S. kudriavzevii*⁶. *S. pastorianus* is thought to originate from the hybrid cross between *S. cerevisiae* and *S. bayanus*⁸ and therefore was not included in this study.

Large chromosomal rearrangements such as reciprocal translocations could act as a post-mating isolating mechanism by inducing the formation of multivalents (an association of more than two homologous chromosomes) during meiosis. The segregation of the centromeres of these multivalents is error-prone, producing aneuploid gametes (containing a number of chromosomes different from the normal haploid complement) and thereby reducing fertility. Two reciprocal translocations have been detected between *S. cerevisiae* and *S. bayanus*, raising the possibility that chromosomal rearrangements are involved in speciation in yeast^{9,10}. To detect reciprocal translocations, we separated chromosomes by clamped homogeneous electric field (CHEF) electrophoresis, transferred them to a nylon membrane and hybridized them with single-gene probes from *S. cerevisiae*. In all six species, the chromosome number is 16 (refs 11, 12) (Fig. 1). We used at least three unique probes per chromosome: one from close to the centromere and one from near each chromosome end.

If the probe corresponding to the left end of chromosome XI in *S. cerevisiae* hybridized to the chromosomal band from one of the other species that was also identified by probes corresponding to both the centromeric region and the right end of *S. cerevisiae* chromosome IV, this would indicate a translocation of the left arm of chromosome XI onto chromosome IV. The reciprocal product is similarly identified. Taking the karyotype of *S. cerevisiae* as a reference, this approach allowed us to detect one non-reciprocal and nine reciprocal translocations involving 13 of the chromosomes (Fig. 1, Table 1). In *S. paradoxus* and *S. kudriavzevii* the karyotypes are colinear with that of *S. cerevisiae* (that is, they show no detectable translocations); therefore, chromosomal rearrangements are not a requirement for speciation in yeast.

We identified four reciprocal translocations involving eight chromosomes in *S. cariocanus*, and one non-reciprocal and three reciprocal translocations in *S. bayanus*. One reciprocal translocation was detected in *S. mikatae* in one isolate and an additional

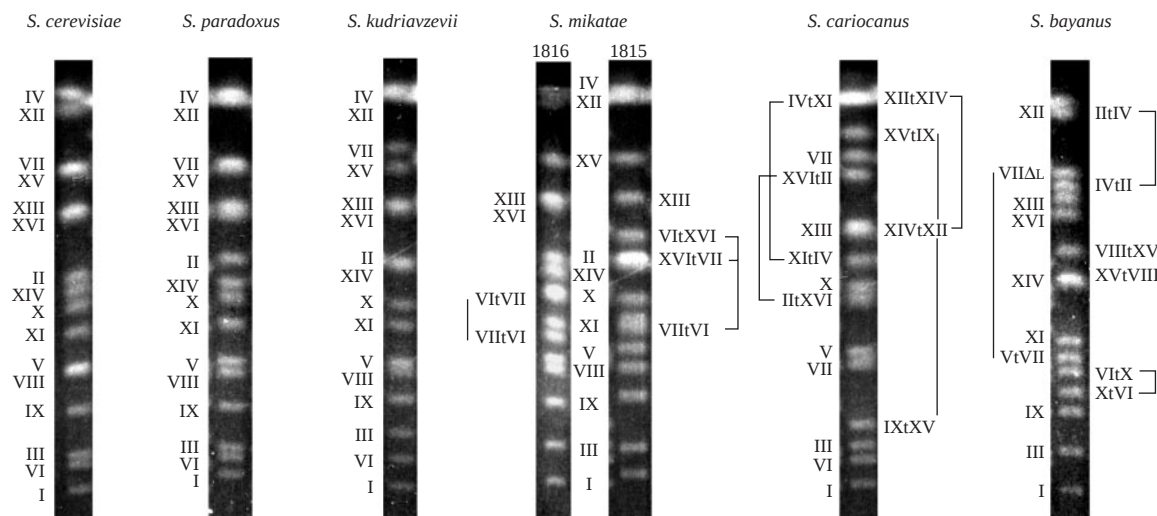


Figure 1 Electrophoretic karyotypes of the *Saccharomyces* 'sensu stricto' species. Strains presented here are *S. cerevisiae* Y55, *S. paradoxus* N17, *S. kudriavzevii* IFO 1802, *S. mikatae* IFO 1816 and IFO 1815, *S. cariocanus* IMUFRJ 50816 and *S. bayanus* CBS 7001. Chromosomes are labelled from I to XVI according to the *S. cerevisiae* nomenclature. Bands showing double intensities correspond to doublets where two non-

homologous chromosomes run at the same position. A triplet involving chromosomes II, XIV and XVI t VII is present in the *S. mikatae* IFO 1815 karyotype. Pairs of chromosomes involved in a reciprocal translocation are connected. In *S. bayanus*, the non-reciprocal translocation event is depicted as V t VII connected to VII Δ (deletion of the left arm of chromosome VII).

translocation was found in the second. In the second isolate, the second translocation involved one of the original translocated chromosomes (Fig. 1, Table 1). Thus, translocations appear to be clustered in a few lineages and each translocation is species-specific, as none of them was found in common amongst the six yeasts studied.

These clustered rearrangements make construction of a chromosomal phylogeny difficult. We constructed phylogenetic trees based on the comparison of the sequences of the internal transcribed spacer regions of the ribosomal RNA genes (ITS-1 and ITS-2). These regions are more variable than the 18S and 26S regions and can be used for comparisons of very closely related species¹³. Using *S. castellii* as the outgroup, two separate phylogenetic taxa can be defined: the 'cerevisiae' (including *S. cerevisiae*, *S. paradoxus*, *S. cariocanus*, *S. mikatae* and *S. kudriavzevii*), and the 'bayanus' (including only *S. bayanus*) (Fig. 2a). Within the *cerevisiae* taxon, the three species sharing colinear karyotypes (*S. cerevisiae*, *S. paradoxus* and *S. kudriavzevii*) lie in well separated branches, indicating that the last common ancestor of this taxon probably had a karyotype colinear with that of the present-day *S. cerevisiae*. *S. cariocanus* (harbouring four reciprocal translocations) is most closely related to *S. paradoxus* (showing no rearrangements).

This tree topology was further supported by using the sequence of the mitochondrial gene *COX2* to construct a second tree. The resulting topology is completely identical to that described above (Fig. 2b). However, the branch lengths within a tree show tips that differ substantially in their distance from the root, and the relative lengths of branches are different between the two trees, indicating that these sequences may not behave like a molecular clock. To test whether the observed clustering of rearrangements is different from a random distribution, we assume *a priori* the tree topology supported by both the nuclear and the mitochondrial data. This topology is used in a tree where branch lengths separating two species from their common ancestor are equalized (analogous to a linearized tree¹⁴). In this tree, the branch lengths represent the time of divergence between species. In Fig. 2c, when the branch length x tends to zero, the branch length y tends to z . This is the maximum divergence time of *S. cariocanus* from *S. paradoxus* and *S. cerevisiae*. Given that the karyotypes of *S. cerevisiae* and *S. paradoxus* are colinear, the probability of accumulating four reciprocal translocations in the *S. cariocanus* lineage is less than

$$\lim_{y \rightarrow z} \left[\left(\frac{y}{2z + y} \right)^4 \right] = \left(\frac{1}{3} \right)^4 = 0.012$$

(assuming a constant rate of formation of rearrangements). This value of P is the upper limit and the probability decreases quickly as the branch length x increases. Similarly, in the 'cerevisiae' taxon, the probability of having all six reciprocal translocations accumulated only in the branches leading to *S. cariocanus* and *S. mikatae* is less than $(2/5)^6 = 0.004$.

Without any *a priori* identification of the species containing the rearrangements, more conservative probabilities can be calculated. Thus, the probability of there being four translocations in just one of the three species (*S. cerevisiae*, *S. paradoxus* and *S. cariocanus*) would be less than $3 \times (1/3)^4 = 0.037$, which is still significant. Similarly, the probability of having the six translocations accumulated in only two out of the five species of the *cerevisiae* taxon would be less than $10 \times (2/5)^6 + 5 \times (1/5)^6 = 0.04032$ (as there are ten possible trees in which the translocations are found in only two of the five species, and five trees in which all of the translocations are found in only one species). These significant probabilities are highly conservative. The actual probabilities are much lower, given that we know the topology and do not believe that the species comprise a star phylogeny. These low probabilities support the idea that the translocations are not randomly distributed; therefore, the rate of formation of chromosomal rearrangements in *Saccharomyces* is not constant. This indicates that bursts of translocations might have occurred at some point in the evolutionary history of some of these species.

Given the high level of relatedness between *S. paradoxus* and *S. cariocanus* (Fig. 2), we tested the possible role of reciprocal translocations in driving speciation between them. Taking 50% as the maximum spore death due to a single reciprocal translocation¹⁵, a cross between two strains sharing four reciprocal translocations should give a minimum viability of $100/2^4 = 6.25\%$. We found a spore viability of 0.66%, consistent with previous findings¹¹ and similar to that obtained from a cross between two species with colinear genomes (such as *S. cerevisiae* and *S. paradoxus*, or *S. cerevisiae* and *S. mikatae*^{3,12}). This finding shows that the present reproductive isolation between *S. paradoxus* and *S. cariocanus* is not due only to the reciprocal translocations, although they might have contributed to it when they occurred. Alternatively, the translocations could have accumulated after the speciation process, as it is the case for the second translocation characterized in *S. mikatae* IFO 1815.

We mapped the translocation breakpoints by first comparing the size of each pair of translocated chromosomes with that of the corresponding chromosomes of *S. cerevisiae*. We then probed

Table 1 Chromosomal translocations in the *S. cerevisiae* 'sensu stricto' group

Species	Translocation	Breakpoint mapping interval	Length of the translocated DNA (kb)	Possible recombining sequences
<i>S. cariocanus</i> IMUFRJ 50816	XVI t II _R	YBR115c–YBR119w	336 (±3)	TEF2 (YBR118w)
	II t XVI _R	YPR075c–YPR082c	248 (±4)	TEF1 (YPR080w)
	XV t IX _L	YIL014w–YIL015w	324 (±3)	Ty1 and Ty3 LTRs
	IX t XV _L	YOL054w–YOL055c	226 (±2)	Ty1 and Ty3 LTRs
	XIV t XII _L	YLL023c–YLL026w	94 (±3.5)	Ty1 LTR
	XII t XIV _L	YNL284c–YNL286w	89 (±3.5)	Ty1
	XI t IV _L	YDL006w–YDL008w	437 (±3)	Ty1 and Ty4 LTRs
	IV t XI _L	YKL059c–YKL065c	322 (±5)	
<i>S. mikatae</i> IFO 1816	VII t VI _R	YFR006w–YFR009w	122 (±2.5)	Ty1 LTRs
	VI t VII _R	YGR021w–YGR026w	564 (±2.2)	tRNA – Asp*
	VII t VI _R	YFR006w–YFR009w	122 (±2.5)	Ty1 LTRs
	(VI t VII _R) t XVI _L	YPL103c–YPL116w	345 (±13)	tRNA – Met*
<i>S. mikatae</i> IFO 1815	XVI t VII _R	YGR188c–YGR189w	219 (±0.8)	Ty1 LTR
	XV t VIII _R †	YHR014w–YHR015w	431 (±0.8)	Ty1 LTRs
	VIII t XV _R †	YOR016c–YOR057w	695 (±35.5)	
	X t VI _R	YFR031c–YFR037c	54 (±4)	tRNA – Lys*
<i>S. bayanus</i> CBS 7001	VI t X _L	YJL044c–YJL047c	355 (±4)	Ty1 LTRs
	V t VII _L	YGL228w–YGL256w	42 (±26)	
	IV t II _R ‡	YBR030w–YBR032w	513	RPL2A (YBR031w)
	II t IV _R ‡	YDR011w–YDR013w	1,060	RPL2B (YDR012w)

XII t XIV_L means a translocation of the left chromosomal arm of chromosome XIV onto chromosome XII.

*See text

†Previously noted¹⁰

‡Previously characterized⁹

electrophoretic karyotypes with labelled open reading frame (ORF) DNAs that fell within the predicted intervals containing the breakpoints. We mapped the translocation breakpoints within relatively short chromosomal intervals for *S. cariocanus* (less than 10 kilobases (kb)) and *S. mikatae* (less than 5 kb (26 kb for the second translocation); Table 1). In *S. cariocanus*, the reciprocal translocation between chromosomes XVI and II is probably due to an ectopic recombination event between the duplicated *TEF1* and *TEF2* genes, which are 99% identical at the DNA level in the sequenced *S. cerevisiae* genome. These genes lie in one of the 55 cluster homology regions (CHRs) found in the yeast genome (block 7)⁴. CHRs are thought to be traces of the whole-genome duplication that occurred in the ancestral yeast⁴.

Similarly, it has been shown that, in *S. bayanus*, the reciprocal translocation between chromosomes IV and II resulted from an ectopic recombination event between duplicated genes belonging to CHR block 3 (ref. 9). In the other cases, Ty elements (full-length transposons or LTRs) have been found within 9 of the 15 remaining breakpoint intervals (Table 1). In addition, transfer RNA sequences are present within three intervals where no repeated elements have

been characterized (Table 1). As there is a close association between Ty elements and tRNA genes¹⁶ (around 90% of the Ty insertions belonging to the four classes Ty1–Ty4 are found near the tRNAs), these sequences could have been used for Ty insertion. Altogether, we found 20 tRNA genes within the 231.6 kb of the breakpoint intervals.

Given that the tRNAs appear to be randomly distributed in the genome¹⁶, the binomial probability of observing 20 or more of the 274 tRNAs in 231.6 kb out of 12,500 kb of non-rDNA genomic sequences is 3×10^{-7} . The 331 Ty and LTR elements are not randomly distributed in the genome but rather show a clumped distribution. We define a Ty cluster as a region between two consecutive ORFs containing at least one Ty or LTR sequence. The 331 elements are then clustered in 193 locations in the genome. The multinomial probability of observing Ty clusters within 9 or more of the 17 breakpoint intervals of 13.6 kb on average (231.6/17) when the average occurrence in the genome is one Ty cluster in every 64 kb (12,500/193) is 5×10^{-3} . Using the actual lengths of the breakpoint intervals, the multinomial probability of observing Ty clusters in six or more of the ten intervals corresponding to the four translocations in *S. cariocanus* and the translocation common to both isolates of *S. mikatae* (leaving out *S. bayanus* because of a probable high level of polymorphism with *S. cerevisiae* in the localization of the Ty elements) is 2×10^{-5} . These low *P* values show that breakpoints occur preferentially in regions with tRNAs and Ty elements, implying that chromosomal translocations are likely to involve ectopic recombination between these dispersed repeated sequences.

Ectopic recombination between Ty sequences occurs at low frequency under laboratory conditions^{17,18}. However, Ty-mediated reciprocal translocations are found in industrial strains^{19,20}. This might be related to the recent adaptation of these yeasts to a new environment, as sequence alterations involving Ty elements and chromosomal rearrangements have been correlated with adaptive changes during long-term cultivation experiments in nutrient-limited media^{21,22}.

Little is known about the mechanisms that control genome stability, but ectopic interactions are stimulated in a mismatch-repair-defective background²³. In addition, a role for sequence divergence and mismatch repair in the establishment of postzygotic species barriers in the *Saccharomyces sensu stricto* group has been proposed³. Perhaps, during adaptation to a new environment, a mutator strain defective in a system controlling the level of ectopic recombination could be selected, as has been observed in the experimental evolution of *Escherichia coli*²⁴. This could account for the apparent bursts of translocations observed during yeast genome evolution but would not in itself necessarily lead to the emergence of new species. □

Methods

Yeast strains, genomic DNA preparation and hybridization

Strains were grown overnight, at 30 °C, in 5 ml of liquid yeast extract peptone dextrose²⁵ [Author: please define]. Genomic DNA for CHEF analysis was prepared and gels were run as described²⁶. The *S. cerevisiae* ORF DNA was supplied by Research Genetics Inc. and directly labelled using the Gene Images kit (Amersham). Southern transfers were performed onto Hybond N⁺ membrane (Amersham), probed with the fluorescein-labelled ORF DNA and detected according to the Gene Images detection module (Amersham). We used 53 ORF probes to detect the translocations and 113 to map the breakpoints.

DNA sequencing and analysis

The entire ITS regions were amplified and sequenced as described¹³, except that primer P3490 (5'-CCGCACGCGCTACACTGA-3'; positions 1,454–1,473 of the *S. cerevisiae* 18S rRNA gene) was used in place of primer pITS1. The primers used for COX2 amplification were 5'-GGTATTTTGAATTACATGA-3' and 5'-ATT-TATTGTTTCGTTTAATCA-3', and the amplification parameters were: 94 °C for 2.5 min followed by two cycles of 94 °C for 2 min, 45 °C for 1.5 min, 72 °C for 1.5 min, followed by 33 cycles of 92 °C for 1.5 min, 45 °C for 1.5 min, 72 °C for 1.5 min, with a final extension step for 5 min at 72 °C. The ITS and COX2 sequences were aligned using the multiple-

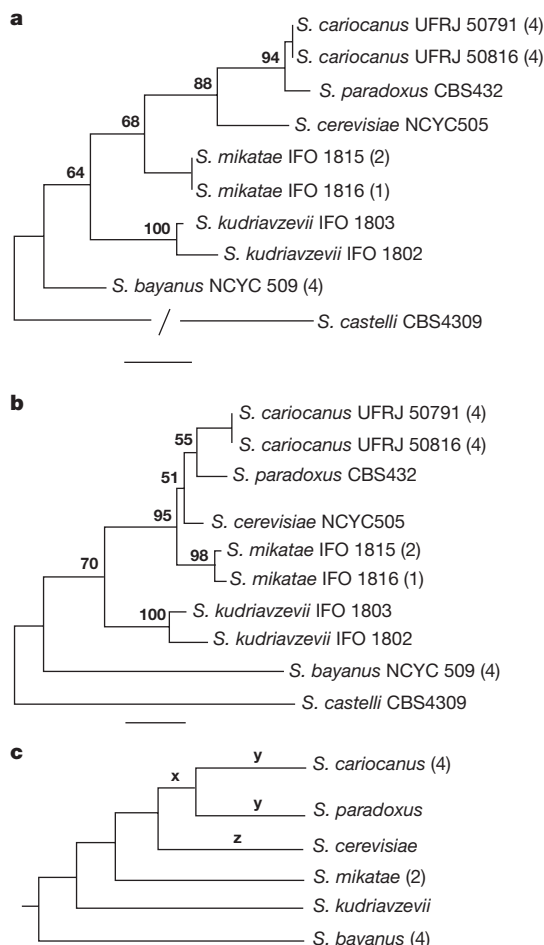


Figure 2 Phylogenetic relationships between the *Saccharomyces cerevisiae* 'sensu stricto' species. **a**, Tree based on the ITS1 sequences. **b**, Tree based on the COX2 sequences. **c**, Linearized tree. *S. castellii* CBS4309 was chosen as the outgroup. The scale bars represent 1 base substitution per 100 nucleotide positions and the bold numbers refer to bootstrap proportions. The number of chromosomal translocations is indicated in parentheses. The ITS1 sequences have been also used for the formal description of the *Saccharomyces sensu stricto* complex as the 18S rRNA sequences were too similar to provide any resolution within this group of species⁶. The topology of the ITS tree is consistent with the ITS1 and 2 trees of the sensu stricto species analysed previously³⁰.

sequence alignment program PILEUP²⁷ and manual adjustment. Phylogenetic analyses were performed using the PHYLIP phylogeny inference package version 3.572. Distance matrices were generated using the DNADIST program with the Jukes–Cantor distance measure. Using *S. castellii* as the outgroup, rooted phylogenetic trees, based on the ITS1 and the COX2 sequences, were constructed by using the neighbour-joining method²⁸ and the NEIGHBOR program. We assessed the stability of the individual branches using the bootstrap method²⁹ with the SEQBOOT, DNADIST, NEIGHBOR and CONSENSE programs.

Received 17 February; accepted 13 March 2000.

- White, M. *Modes of Speciation* (Freeman, San Francisco, 1978).
- Naumov, G. Genetic basis for classification and identification of the ascomycetous yeasts. *Stud. Mycol.* **30**, 469–475 (1987).
- Hunter, N., Chambers, S. R., Louis, E. J. & Borts, R. H. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* **15**, 1726–1733 (1996).
- Wolfe, K. H. & Shields, D. C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708–713 (1997).
- Seoighe, C. & Wolfe, K. H. Extent of genomic rearrangement after genome duplication in yeast. *Proc. Natl Acad. Sci. USA* **95**, 4447–4452 (1998).
- Naumov, G., James, S. A., Louis, E. J., Naumova, E. S. & Roberts, I. N. Three new species in the *Saccharomyces sensu stricto* complex: *Saccharomyces cariocanus*, *Saccharomyces kudriavzevii* and *Saccharomyces mikatae*. *Int. J. Syst. Bacteriol.* (in the press).
- James, S. A., Cai, J., Roberts, I. N. & Collins, M. D. A phylogenetic analysis of the genus *Saccharomyces* based on 18S rDNA gene sequences: description of *Saccharomyces kunashirensis* sp. nov. and *Saccharomyces martiniae* sp. nov. *Int. J. Syst. Bacteriol.* **47**, 453–460 (1997).
- Tamai, Y., Momma, T., Yoshimoto, H. & Kaneko, Y. Co-existence of 2 types of chromosome in the bottom fermenting yeast *Saccharomyces pastorianus*. *Yeast* **10**, 923–933 (1998).
- Ryu, S. L., Murooka, Y. & Kaneko, Y. Reciprocal translocation at duplicated RPL2 loci might cause speciation of *Saccharomyces bayanus* and *Saccharomyces cerevisiae*. *Curr. Genet.* **33**, 345–351 (1998).
- Ryu, S. L., Murooka, Y. & Kaneko, Y. Genomic reorganization between two sibling yeast species, *Saccharomyces bayanus* and *Saccharomyces cerevisiae*. *Yeast* **12**, 757–764 (1996).
- Naumov, G. I., Naumova, E. S., Hagler, A. N., Mendoncahagler, L. C. & Louis, E. J. A new genetically isolated population of the *Saccharomyces sensu stricto* complex from Brazil. *Antonie Van Leeuwenhoek* **67**, 351–355 (1995).
- Naumov, G. I., Naumova, E. S. & Louis, E. J. Two new genetically isolated populations of the *Saccharomyces sensu stricto* complex from Japan. *J. Gen. Appl. Microbiol.* **41**, 499–505 (1995).
- James, S. A., Collins, M. D. & Roberts, I. N. Use of an rRNA internal transcribed spacer region to distinguish phylogenetically closely related species of the genera *Zygosaccharomyces* and *Torulaspora*. *Int. J. Syst. Bacteriol.* **46**, 189–194 (1996).
- Takezaki, N., Rzhetsky, A. & Nei, M. Phylogenetic test of the molecular clock and linearized tree. *Mol. Biol. Evol.* **12**, 823–833 (1995).
- Loidl, J., Jin, Q. W. & Jantsch, M. Meiotic pairing and segregation of translocation quadrivalents in yeast. *Chromosoma* **107**, 247–254 (1998).
- Hani, J. & Feldmann, H. tRNA genes and retroelements in the yeast genome. *Nucleic Acids Res.* **26**, 689–696 (1998).
- Boeke, J. D., Eichinger, D. J. & Natsoulis, G. Doubling Ty1 element copy number in *Saccharomyces cerevisiae*: host genome stability and phenotypic effects. *Genetics* **129**, 1043–1052 (1991).
- Kupiec, M. & Petes, T. D. Meiotic recombination between repeated transposable elements in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **8**, 2942–2954 (1988).
- Rachidi, N., Barre, P. & Blondin, B. Multiple Ty-mediated chromosomal translocations lead to karyotype changes in a wine strain of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **261**, 841–850 (1999).
- Codon, A. C., Benitez, T. & Korhola, M. Chromosomal reorganization during meiosis of *Saccharomyces cerevisiae* baker's yeasts. *Curr. Genet.* **32**, 247–259 (1997).
- Adams, J. & Oeller, P. W. Structure of evolving populations of *Saccharomyces cerevisiae* adaptive changes are frequently associated with sequence alterations involving mobile elements belonging to the Ty family. *Proc. Natl Acad. Sci. USA* **83**, 7124–7127 (1986).
- Adams, J., Puskasroza, S., Simlar, J. & Wilke, C. M. Adaptation and major chromosomal changes in populations of *Saccharomyces cerevisiae*. *Curr. Genet.* **22**, 13–19 (1992).
- Bailis, A. M. & Rothstein, R. A defect in mismatch repair in *Saccharomyces cerevisiae* stimulates ectopic recombination between homologous genes by an excision repair dependent process. *Genetics* **126**, 535–547 (1990).
- Snigowski, P., Gerrish, P. & Lenski, R. Evolution of high mutation rates in experimental populations of *Escherichia coli*. *Nature* **387**, 703–705 (1997).
- Sherman, F., Fink, G. & Hicks, J. *Laboratory Course Manual for Methods in Yeast Genetics* (Cold Spring Harbour Laboratory, Cold Spring Harbour, 1986).
- Louis, E. J. in *Methods in Microbiology; Yeast Gene Analysis* (eds Truite, M. F. & Brown, A. J. P.) 15–32 (Academic, London, 1998).
- Feng, D. F. & Doolittle, R. F. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **35**, 351–360 (1987).
- Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
- Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791 (1985).
- Goddard, M. R. & Burt, A. Recurrent invasion and extinction of a selfish gene. *Proc. Natl Acad. Sci. USA* **96**, 13880–13885 (1999).

Acknowledgements

We thank R. Borts and S. Chambers for helpful comments on the manuscript. This work was supported by the Biotechnology and Biological Sciences Research Council and the Wellcome Trust.

Correspondence and requests for materials should be addressed to E.J.L. (e-mail: elouis@molbiol.ox.ac.uk).

Synaptic activity at calcium-permeable AMPA receptors induces a switch in receptor subtype

Si-Qiong June Liu & Stuart G. Cull-Candy

Department of Pharmacology, University College London, Gower Street, London WC1E 6BT, UK

Activity-dependent change in the efficacy of transmission is a basic feature of many excitatory synapses in the central nervous system. The best understood postsynaptic modification involves a change in responsiveness of AMPAR (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptor)-mediated currents following activation of NMDA (N-methyl-D-aspartate) receptors^{1,2} or Ca²⁺-permeable AMPARs^{3–6}. This process is thought to involve alteration in the number and phosphorylation state of postsynaptic AMPARs². Here we describe a new form of synaptic plasticity—a rapid and lasting change in the subunit composition and Ca²⁺ permeability of AMPARs at cerebellar stellate cell synapses following synaptic activity. AMPARs lacking the edited GluR2 subunit not only exhibit high Ca²⁺ permeability⁷ but also are blocked by intracellular polyamines^{8–11}. These properties have allowed us to follow directly the involvement of GluR2 subunits in synaptic transmission. Repetitive synaptic activation

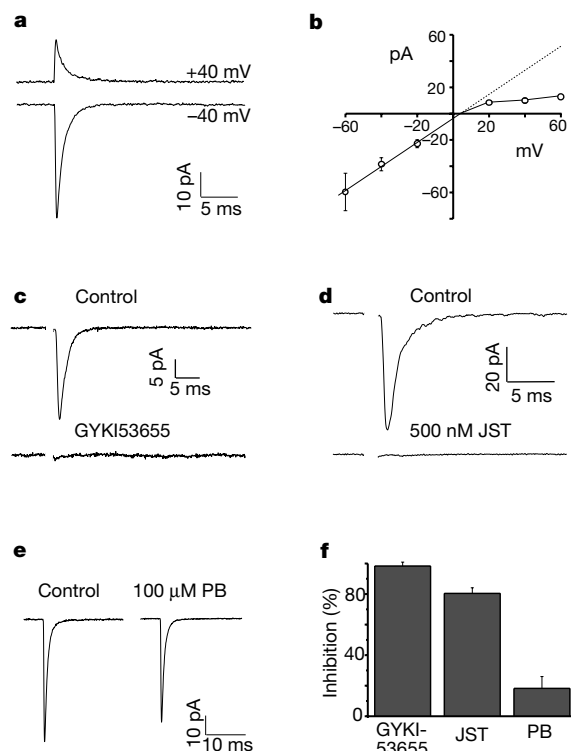


Figure 1 Synaptic currents in stellate cells exhibit properties of GluR2-lacking AMPARs. **a**, Mean sEPSC at +40 and –40 mV, showing reduced current at positive potentials (rectification occurred only when spermine was included in pipette). **b**, Rectifying *I*–*V* relationship of sEPSCs (*n* = 5). Solid line follows the data points; dashed line at positive potentials represents extrapolated fit for EPSCs behaving ohmically. **c**, Evoked EPSCs (averaged traces) were blocked by the selective AMPAR antagonist GYKI53655 (25 μ M). **d**, Joro spider toxin (JST, 500 nM) blocked EPSCs (averaged traces) evoked by threshold stimulation of parallel fibres. **e**, Mean sEPSCs were slightly reduced but not blocked by pentobarbital (PB, 100 μ M). **f**, Percentage inhibition of EPSCs by GYKI53655 (*n* = 5), JST (*n* = 5) and PB (*n* = 5; –60 mV). Effect of GYKI53655 and JST was significant (*P* < 0.01).

of Ca^{2+} -permeable AMPARs causes a rapid reduction in Ca^{2+} permeability and a change in the amplitude of excitatory postsynaptic currents, owing to the incorporation of GluR2-containing AMPARs. Our experiments show that activity-induced Ca^{2+} influx through GluR2-lacking AMPARs controls the targeting of GluR2-containing AMPARs, implying the presence of a self-regulating mechanism.

Minimally evoked parallel fibre input to cerebellar stellate cells is mediated solely by non-NMDA receptors¹². We found that the selective antagonist GYKI53655 (10–25 μM) reduced the amplitude of evoked excitatory postsynaptic currents (eEPSCs) by $98 \pm 3\%$ ($n = 5$; Fig. 1c, f), indicating that these currents are mediated by AMPARs. We have used a number of approaches to determine whether EPSCs and agonist-evoked currents are mediated by the Ca^{2+} -permeable variety of AMPARs in these cells. Inclusion of spermine in the pipette solution is known to confer voltage-dependent block on AMPARs lacking GluR2 subunits, which differ from other subunits in being edited in their pore-lining region. This produces a characteristic inwardly rectifying $I-V$ relationship^{8–11}. The amplitude of spontaneous EPSCs (sEPSCs) (Fig. 1a, b) at the parallel fibre input was reduced at positive membrane potentials. This rectifying $I-V$ relationship implies that the EPSCs are mediated mainly by Ca^{2+} -permeable AMPARs. We have used various other pharmacological approaches to determine that most synaptic receptors do indeed lack GluR2 subunits.

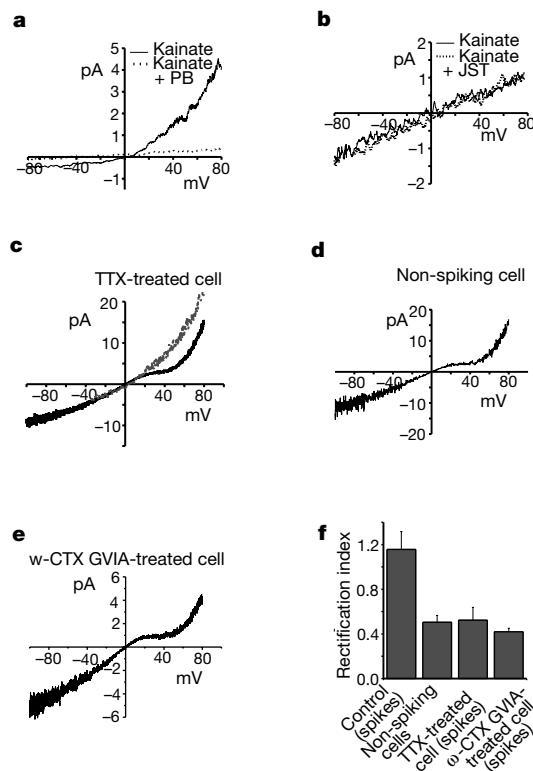


Figure 2 Extrasynaptic patches exhibit $I-V$ plots (outwardly rectifying in **a**; linear in **b**), indicative of GluR2-containing AMPARs. These convert to GluR2-lacking AMPARs in the absence of action potentials, or absence of Ca^{2+} influx through N-type channels.

a, Kainate-evoked current (solid line) was blocked by pentobarbital (PB, 100 μM ; dashed line). **b**, Joro spider toxin (500 nM, thick line) did not inhibit the current. **c**, Inwardly rectifying $I-V$ plot after treatment with 500 nM TTX for more than 2 h before recording (thick line); compare this at positive potentials with an averaged control $I-V$ plot (thin line). TTX absent during recording. **d**, Inwardly rectifying $I-V$ plot in a patch from a cell that did not fire spontaneously action potentials. **e**, Inwardly rectifying $I-V$ plot after treatment with 500 nM ω -conotoxin (CTX) GVIA for more than 2 h before recording. **f**, Rectification index was significantly reduced in patches from non-spiking cells ($n = 5$), and cells treated with TTX ($n = 6$) or CTX GVIA ($n = 7$), when compared with control cells ($n = 13$) ($P < 0.03$).

Thus, Joro spider toxin (JST), a subunit-specific blocker of the GluR2-lacking AMPARs¹³, reduced the EPSC amplitude by $80 \pm 4\%$ ($n = 5$) at -60 mV (Fig. 1d). On the other hand, pentobarbital, which at the concentration used selectively inhibits the GluR2-containing AMPARs¹⁴, reduced the EPSC size by only $\sim 18\%$ ($n = 5$; Fig. 1e, f).

In contrast to the synaptic currents, agonist-evoked currents in somatic outside-out patches exhibited linear or outwardly rectifying $I-V$ relationships (Fig. 2a, b) that were blocked by GYKI53655 ($90 \pm 15\%$, $n = 5$). This suggests the presence of the Ca^{2+} -impermeable variety of AMPARs in the extrasynaptic membrane. Furthermore, pentobarbital reversibly inhibited these currents by $82 \pm 13\%$ (at -60 mV; $n = 5$; Fig. 2a), while JST gave only a $\sim 10 \pm 15\%$ reduction ($n = 6$; Fig. 2b). These findings indicate that stellate cells express a population of GluR2-containing receptors, but that these occur predominantly extrasynaptically.

To investigate the factors responsible for the selective targeting of GluR2 in these cells, we tested whether synaptic activity could regulate the Ca^{2+} permeability of synaptic AMPARs. After high-frequency synaptic stimulation (300 stimuli at 50 Hz at -60 mV),

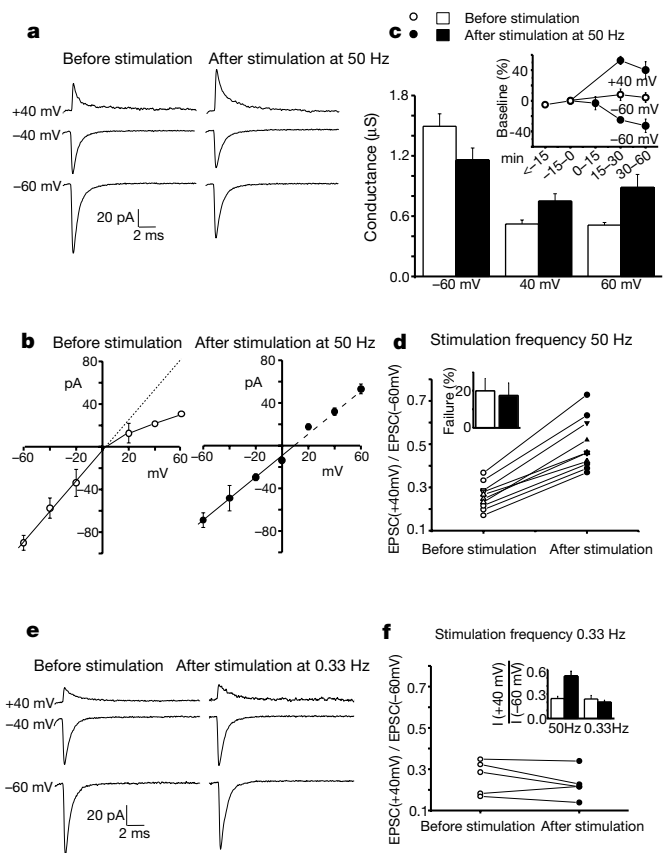


Figure 3 High-frequency synaptic stimulation induced a change in the rectification of $I-V$ relationships of EPSCs. **a**, After high-frequency stimulation (300 stimuli at 50 Hz), EPSC amplitude was reduced at -60 mV, but increased at $+40$ mV. **b**, The inwardly rectifying $I-V$ relationship before stimulation (open circles), became linear after 50-Hz stimulation (filled circles) ($n = 9$). Before stimulation the data at positive potentials fell below the line extrapolated from a linear fit; after stimulation data at positive potentials fell along the extrapolated line. **c**, Averaged EPSC conductance values (-60 , $+40$ and $+60$ mV) changed significantly after stimulation ($P < 0.05$). Inset, stimulation induces a lasting change in EPSC amplitude. **d**, High-frequency stimulation consistently increased the ratio of EPSC amplitudes at $+40$ mV versus -60 mV but did not change failure rate. Cells were stimulated in the presence of either 120 μM D-AP5/10 μM 7-chlorokynurenic acid (at -60 mV inverted triangle; at -90 mV triangle) or 20 μM D-AP5 (at -60 mV circle; same data set as in **b** and **c**). **e**, **f**, Low-frequency stimulation (300 stimuli at 0.33 Hz) did not alter EPSC amplitude or rectification ($n = 5$). Inset, mean ratio of EPSC amplitude at $+40$ versus -60 mV increases by more than twofold after high-frequency stimulation at 50 Hz ($P < 0.0005$, paired t -test) but remains unchanged after stimulation at 0.33 Hz.

the I - V relationship of eEPSCs was changed from predominantly inwardly rectifying (rectification index (RI) = 0.43 ± 0.04 , $n = 6$) to a linear form (RI = 1.04 ± 0.14 , $n = 6$) (see Figs 3a, b and 4f) within ~ 15 – 30 min (Fig. 3c). The EPSC amplitude was significantly decreased at negative potentials (from -93 ± 7 to -65 ± 5 pA at -60 mV; $n = 8$, $P < 0.005$, paired t -test) but increased at positive potentials (from 23 ± 2 to 32 ± 4 pA at $+40$ mV, $n = 6$, $P < 0.02$, paired t -test). The decay time constant was little changed (from 0.76 ± 0.04 ms to 0.89 ± 0.08 ms; $n = 11$, $P < 0.05$). The ratio of EPSC amplitudes at $+40$ mV versus -60 mV increased in all cells examined (without a change in the percentage of failures), regardless of whether cells were held at -60 or -90 mV during stimulation (Fig. 3d). Together, these data indicate an increased expression of the GluR2 subunit at the synapse, which produces a loss of polyamine block and removal of rectification. The reduced EPSC amplitude at -60 mV could reflect changed elementary channel properties as occurs with GluR2-containing recombinant heterooligomers¹⁵. In contrast, stimulation at a low frequency, with the same number of stimuli (300 stimuli at 0.33 Hz), did not change the EPSC amplitude or the rectification (Fig. 3e, f). Thus, high-frequency stimulation triggers the expression of Ca^{2+} -impermeable AMPARs at these synapses. Furthermore, the change was not inhibited by the presence of $120 \mu\text{M}$ D-AP5 or $10 \mu\text{M}$ 7-chloro-

kyurenic acid (Fig. 3d), indicating that NMDARs were not required for this process. If this activity-dependent change in AMPAR expression is a physiological mechanism then it is expected to be continuously occurring *in vivo*. Hence different amounts of rectification would be seen at different synapses. In keeping with this idea, we found that the ratio of EPSC amplitudes at $+40$ to those at -60 mV ranged from 0.116 to 0.394 in control cells ($n = 42$).

We then examined whether the effect of high-frequency synaptic activity could be reproduced by bath application of 1 mM glutamate. After glutamate application, the I - V relationship of the EPSC became near linear (for at least 2–3 h) (Fig. 4d). We also found that bath application of $100 \mu\text{M}$ kainate, which activates non-NMDA receptors but not metabotropic receptors¹⁶, gave rise to EPSCs with linear I - V relationships and an RI of 0.92 ± 0.15 ($n = 5$; Fig. 4f). Thus, direct activation of AMPARs can produce an increase in GluR2-containing receptors at this synapse.

To determine whether Ca^{2+} entry was the trigger necessary for activity-dependent targeting of GluR2, we used three different experimental approaches. First, we recorded EPSCs with BAPTA ($1,2$ -bis(*o*-Aminophenoxy)ethane-*N,N,N',N'*-tetra-acetic acid) (20 mM) included in the pipette to inhibit Ca^{2+} increase in the postsynaptic cell. The high-frequency stimulation no longer altered EPSC amplitude at any of the potentials tested (Fig. 4a,b). Second, cells were held at $+40$ mV during the period of stimulation, to reduce Ca^{2+} influx through synaptic channels. Again, no change in the EPSC amplitude was observed (Fig. 4c). Third, glutamate was applied in the absence of external Ca^{2+} . In these conditions, glutamate failed to produce any change in the inwardly rectifying I - V relationship (Fig. 4e), consistent with a requirement for Ca^{2+} entry. Our experiments using high-frequency synaptic stimulation (see above) in cells clamped at -90 mV indicate that a change in GluR2-containing receptors occurs in the absence of activation of voltage-gated Ca^{2+} channels. These data are consistent with the idea that Ca^{2+} influx through the AMPARs is sufficient to produce an intracellular Ca^{2+} rise that causes a rapid change in subunit composition of synaptic receptors.

Does Ca^{2+} entry through voltage-gated Ca^{2+} channels influence the expression of GluR2? Stellate cells fire spontaneous action potentials in the absence of synaptic inputs (mean firing rate $\sim 12 \text{ Hz}$)¹⁷. We blocked action potential firing (with 500 nM tetrodotoxin (TTX)) for at least two hours, and then examined the agonist-evoked currents in outside-out patches from the soma. The I - V relationship was changed from its usual linear or outwardly rectifying form (Fig. 2a, b) to one that showed strong inward rectification (Fig. 2c), consistent with induction of Ca^{2+} -permeable extrasynaptic receptors. The rectification index of the patches from TTX-treated cells was reduced to 0.5 (from control levels of ~ 1.2 ; see Fig. 2f). Most of the stellate cells fired spontaneous action potentials in the absence of TTX, but some cells in the molecular layer did not (although excitatory and inhibitory synaptic currents were present). Outside-out patches from these cells also exhibited inwardly rectifying I - V plots (Fig. 2d), indicative of Ca^{2+} -permeable receptors. Furthermore, patches from cells treated with ω -conotoxin GVIA exhibited inwardly rectifying I - V relationships (Fig. 2e). These results indicate that local Ca^{2+} influx through N-type Ca^{2+} -channel activity during action potentials can influence the normal synthesis or targeting of GluR2 subunits, accounting for the absence of Ca^{2+} -permeable extrasynaptic receptors.

Earlier work has shown that Ca^{2+} -permeable AMPARs can be differentially localized at some inputs onto hippocampal neurons¹⁸, although the mechanism underlying this differential distribution is unclear. Our experiments indicate that the level of GluR2-containing receptors at the parallel fibre stellate cell synapse, and the soma in these cells, undergoes dynamic changes driven by activity-induced local Ca^{2+} influx. The synaptic AMPARs are mostly Ca^{2+} permeable and therefore lack the edited GluR2 subunit. During high-frequency stimulation, increased Ca^{2+} entry through synaptic AMPARs triggers

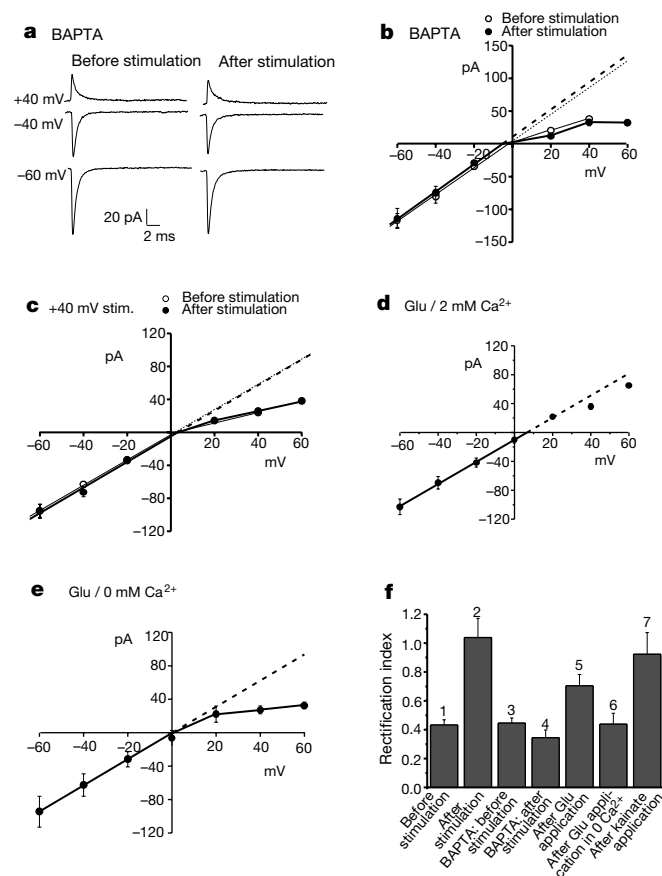


Figure 4 Calcium influx through non-NMDA glutamate receptors is sufficient to trigger activity-dependent change in EPSC rectification. **a**, **b**, Intracellular BAPTA (20 mM) prevented activity-induced (3×100 stimuli at 50 Hz) change in amplitude and inward rectification of EPSCs ($n = 5$). **c**, EPSC amplitude and I - V relationship was unaltered if cells were clamped at $+40$ mV during high-frequency stimulation ($n = 5$). **d**, **e**, EPSCs gave linear I - V relationships in cells ($n = 5$) exposed to bath application of 1 mM glutamate. EPSCs remained inwardly rectifying ($n = 5$) if glutamate was applied in Ca^{2+} -free solution (plus 3 mM MgCl_2). **f**, Summary of rectification index calculated using data in **b**, **d**, **e** and Fig. 3b; I - V relationship obtained after bath application of $100 \mu\text{M}$ kainate and 200 – 500 nM TTX. Differences between 1 and 2, 1 and 5, 1 and 7, 2 and 4, and 5 and 6 were significant ($P < 0.05$).

the insertion of GluR2-containing receptors to the synapse and the removal of GluR2-lacking receptors. This results in a reduced EPSC amplitude at -60 mV, as well as a change in Ca^{2+} permeability. This molecular modification of the receptor would provide self-regulating feedback that further reduces Ca^{2+} entry, and hence limits the level of GluR2-containing receptors. Our finding that GluR2-containing receptors are present in the soma is in keeping with such a mechanism, as local Ca^{2+} influx during action potentials influences GluR2 expression. The presence of GluR2-lacking receptors at synapses, however, may result from attenuation of dendritic invasion by action potentials¹⁹, or from regional differences in Ca^{2+} -channel subtype density²⁰ or Ca^{2+} buffering²¹. Recent studies on CA1 cells have provided evidence for insertion of AMPARs by exocytosis into synaptic regions during long-term potentiation^{22,23}. Furthermore, a specific interaction between NSF (N-ethylmaleimide-sensitive factor), a protein involved in membrane fusion events, and the carboxy terminus of GluR2 subunits has been described^{24–26}. It is therefore feasible that insertion of GluR2-containing receptors, and removal of GluR2-lacking receptors at the synapse, could be mediated by exocytosis and endocytosis, respectively, although redistribution through lateral diffusion cannot be excluded.

Previous observations have shown that neuronal activity can act to regulate EPSC amplitude by changing AMPAR number^{23,27–29} and channel properties³⁰. Our experiments suggest that synaptic activity can directly determine the subunit composition of a synaptic AMPAR by controlling the expression or targeting of edited subunits. This produces a change not only in amplitude of synaptic currents, but also in their calcium permeability, voltage dependence and facilitation properties⁸. Such a mechanism provides a form of plasticity that is self-regulating, and would represent a significant role for edited non-NMDAR subunits in certain postsynaptic modifications. □

Methods

Electrophysiology

Sagittal or coronal cerebellar slices (200 – 250 μm) were cut with a vibrating microslicer (DTK-1000) from the vermis of 18–20-day-old Sprague–Dawley rats in ice-cold solution (in mM: 125 NaCl, 2.5 KCl, 2 CaCl_2 , 1 MgCl_2 , 26 NaHCO_3 , 1.25 NaH_2PO_4 , and 25 glucose, saturated with 95% O_2 /5% CO_2 , pH 7.4). Whole-cell patch-clamp and outside-out patch recordings were made with an Axopatch 200A amplifier (Axon Instruments) in external solution which contained GABA_A and NMDA receptor blockers (20 μM bicuculline methobromide, 100 μM picrotoxin and 20 μM D-AP5) at room temperature. A few stimulation protocols, and all agonist application protocols were carried out in a higher concentration of D-AP5 (120 μM) and 7-chlorokynurenic acid (10 μM).

Recordings were made from visually identified neurons located in the outer two-thirds of the molecular layer. Interneurons were usually identified by their ability to fire spontaneous action potentials in the cell-attached configuration and by the presence of spontaneous EPSCs and inhibitory postsynaptic currents in the whole-cell configuration. Electrode resistances were 3–8 M Ω when filled with internal solution (in mM: 95 CsF, 45 CsCl (or 140 CsMeSO₄), 10 CsHEPES, 10 CsEGTA, 2 NaCl, 2 ATP-Mg, 1 QX314, 5 TEA, 1 CaCl_2 , 0.1 spermine, pH 7.3). Series resistance, whole-cell capacitance and input resistance were 12.6 ± 0.7 M Ω , 4.0 ± 0.3 pF ($n = 33$) and 1.28 ± 0.28 G Ω ($n = 16$), respectively. If series resistance was changed by more than 20%, the experiment was discarded. EPSCs were evoked with a patch-electrode containing external solution, by stimulating in the molecular layer (20–100 μs pulses of 5–30 V at 0.33 Hz) and filtered at 10 kHz. Currents from outside-out patches, in response to voltage ramps (42 mV s^{-1}), were measured before and during the application of 100 μM kainate or 1 mM glutamate + 100 μM cyclothiazide.

Synaptic stimulation protocol and application of agonists

Parallel fibres were stimulated at 50 Hz or 0.33 Hz, while the postsynaptic cell was voltage-clamped at -60 mV (unless otherwise indicated). Only when high-frequency stimulation was successful at generating EPSCs, and there was no change in holding current of the postsynaptic cell, was the data used for further analysis. Stimulation strength and duration were kept constant throughout the experiment. Glutamate and kainate were bath applied for ~ 1 –2 min. Synaptic currents were measured between 15 min and 2 h after synaptic stimulation, or after bath application of agonist.

Analysis

EPSCs were filtered at 4 kHz and digitized at 20 kHz. For the I – V analysis, we rejected events that did not have a smooth rise and decay phase. The average at each holding potential was constructed by aligning each event on its point of fastest rise (typically average of 20–40 EPSCs) using N version 4.0 (written by S. Traynelis, Emory University). The mean EPSC amplitudes at negative potentials were fitted by a linear regression. If

EPSC amplitude at positive potentials fell below the extrapolated line it was considered an inwardly rectifying I – V relationship. The RI of the I – V relationship was defined as the ratio of the current amplitude at $+40$ mV to the predicted linear value at $+40$ mV (extrapolated from linear fitting of the currents at the negative potentials). The inhibition of the EPSC by GYKI53655 and JST was calculated by integrating the average EPSC before and after drug addition and the average was obtained by aligning all the traces to their stimulus artefact.

The I – V plots of agonist-evoked currents in outside-out patches were obtained by subtracting leak current from the current in the presence of kainate or glutamate plus cyclothiazide. The rectification index was defined as the ratio of current amplitude at $+40$ mV to the predicted linear value at $+40$ mV (extrapolated from linear fit of the current between -20 mV and $+10$ mV, as some patches had outwardly rectifying currents). All values are expressed as mean $\pm$ s.e.m., except when they are smaller than the size of the symbol. Statistical significance was assessed by the two-tailed Student's t -test.

Received 30 November 1999; accepted 20 March 2000.

- Bliss, T. V. P. & Collingridge, G. L. A synaptic model of memory—long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
- Malenka, R. C. & Nicoll, R. A. Long-term potentiation—a decade of progress? *Science* **285**, 1870–1874 (1999).
- Gu, J. G., Albuquerque, C., Lee, C. J. & MacDermott, A. B. Synaptic strengthening through activation of Ca^{2+} -permeable AMPA receptors. *Nature* **381**, 793–796 (1996).
- Mahanty, N. K. & Sah, P. Calcium-permeable AMPA receptors mediated long-term potentiation in interneurons in the amygdala. *Nature* **394**, 683–687 (1998).
- Feldmeyer, D. et al. Neurological dysfunctions in mice expressing different levels of the Q/R site-unedited AMPAR subunit GluR-B. *Nature Neurosci.* **2**, 57–64 (1999).
- Laezza, F., Doherty, J. J. & Dingledine, R. Long term depression in Hippocampal interneurons: joint requirement for pre- and postsynaptic events. *Science* **285**, 1411–1414 (1999).
- Geiger, J. R. P. et al. Relative abundance of subunit mRNAs determines gating and Ca^{2+} permeability of AMPA receptors in principal neurons and interneurons in rat CNS. *Neuron* **15**, 193–204 (1995).
- Rozov, A. & Burnashev, N. Polyamine-dependent facilitation of postsynaptic AMPA receptors counteracts paired-pulse depression. *Nature* **401**, 594–598 (1999).
- Bowie, D. & Mayer, M. Inward rectification of both AMPA and kainate subtype glutamate receptors generated by polyamine-mediated ion-channel block. *Neuron* **15**, 453–462 (1995).
- Kamboj, S. K., Swanson, G. T. & Cull-Candy, S. G. Intracellular spermine confers rectification on rat calcium-permeable AMPA and kainate receptors. *J. Physiol.* **486**, 297–303 (1995).
- Koh, D.-S., Burnashev, N. & Jonas, P. Block of native Ca^{2+} -permeable AMPA receptors in rat brain by intracellular polyamines generates double rectification. *J. Physiol.* **486**, 305–312 (1995).
- Clark, B. A. & Cull-Candy, S. Frequency-dependent activation of NMDA receptors at an 'AMPA receptor only' synapse in the rat cerebellum. *J. Physiol.* **518P**, 156P (1999).
- Blaschke, M. et al. A single amino acid determined the subunit-specific spider toxin block of α -amino-3-hydroxy-5-methylisoxazole-4-propionate/kainate receptor channels. *Proc. Natl Acad. Sci. USA* **90**, 6528–6532 (1993).
- Yamkura, T., Sakimura, K., Mishina, M. & Shimoji, K. The sensitivity of AMPA-selective glutamate receptor channels to pentobarbital is determined by a single amino acid residue of the $\alpha 2$ subunit. *FEBS Lett.* **374**, 412–414 (1995).
- Swanson, G. T., Kamboj, S. K. & Cull-Candy, S. G. Single-channel properties of recombinant AMPA receptors depend on RNA editing, splice variation, and subunit composition. *J. Neurosci.* **17**, 58–69 (1997).
- Masu, M., Tanabe, Y., Tsuchida, K., Shigemoto, R. & Nakanishi, S. Sequence and expression of a metabotropic glutamate receptor. *Nature* **349**, 760–765 (1991).
- Hausser, M. & Clark, B. A. Tonic synaptic inhibition modulates neuronal output pattern and spatiotemporal synaptic integration. *Neuron* **19**, 665–678 (1997).
- Toth, K. & McBain, C. J. Afferent-specific innervation of two distinct AMPA receptor subtypes on single hippocampal interneurons. *Nature Neurosci.* **1**, 572–578 (1998).
- Spruston, N., Schiller, Y., Stuart, G. & Sakmann, B. Activity-dependent action potential invasion and calcium influx into hippocampal CA1 dendrites. *Science* **268**, 297–300 (1995).
- Magee, J., Hoffman, D., Colbert, C. & Johnston, D. Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons. *Annu. Rev. Physiol.* **60**, 327–346 (1998).
- Seidenbecher, C. I. et al. Calbindin, a novel neuronal calcium-binding protein confined to the somato-dendritic compartment. *J. Biol. Chem.* **273**, 21324–21331 (1998).
- Lledo, P.-M., Zhang, X., Sudhof, T. C., Malenka, R. C. & Nicoll, R. A. Postsynaptic membrane fusion and long-term potentiation. *Science* **279**, 399–403 (1998).
- Shi, S.-H. et al. Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science* **284**, 1811–1816 (1999).
- Song, I. et al. Interaction of the N-ethylmaleimide-sensitive factor with AMPA receptors. *Neuron* **21**, 393–400 (1998).
- Osten, P. et al. The AMPA receptor GluR2 C terminus can mediate a reversible, ATP-dependent interaction with NSF and α - and β -SNAPs. *Neuron* **21**, 99–110 (1998).
- Nishimune, A. et al. NSF binding to GluR2 regulates synaptic transmission. *Neuron* **21**, 87–97 (1998).
- Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C. & Nelson, S. B. Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* **391**, 892–896 (1998).
- O'Brien, R. J. et al. Activity-dependent modulation of synaptic AMPA receptor accumulation. *Neuron* **21**, 1067–1078 (1998).
- Carroll, R. C., Lissin, D. V., von Zastrow, M., Nicoll, R. A. & Malenka, R. C. Rapid redistribution of glutamate receptors contributes to long-term depression in hippocampal cultures. *Nature Neurosci.* **2**, 454–460 (1999).
- Benke, T. A., Luthi, A., Issac, J. T. & Collingridge, G. L. Modulation of AMPA receptor unitary conductance by synaptic activity. *Nature* **393**, 793–797 (1998).

Acknowledgements

We thank S. Brickley, B. Clark, M. Farrant, C. Misra, S. Traynelis and M. Whim for helpful advice, discussion and comments on the manuscript, and J. Thomas for technical help.

This work was supported by the Wellcome Trust. S.-Q.L. was in receipt of a Wellcome Trust Travelling Fellowship.

Correspondence and requests for materials should be addressed to S.G.C.-C. (e-mail: s.cull-candy@ucl.ac.uk).

Vagus nerve stimulation attenuates the systemic inflammatory response to endotoxin

Lyudmila V. Borovikova*, Svetlana Ivanova*, Minghuang Zhang*, Huan Yang*, Galina I. Botchkina*, Linda R. Watkins†, Haichao Wang‡, Naji Abumrad§, John W. Eaton¶ & Kevin J. Tracey§

* The Picower Institute for Medical Research, Manhasset, New York 11030, USA

† Department of Psychology, University of Colorado, Boulder, Colorado 80309, USA

‡ Department of Emergency Medicine and § Department of Surgery, North Shore University Hospital, Manhasset, New York 11030, USA

¶ Department of Pediatrics, Baylor College of Medicine, Houston, Texas 77030, USA

Vertebrates achieve internal homeostasis during infection or injury by balancing the activities of proinflammatory and anti-inflammatory pathways. Endotoxin (lipopolysaccharide), produced by all gram-negative bacteria, activates macrophages to release cytokines that are potentially lethal^{1–4}. The central nervous system regulates systemic inflammatory responses to endotoxin through humoral mechanisms^{5–8}. Activation of afferent vagus nerve fibres by endotoxin or cytokines stimulates hypothalamic–

pituitary–adrenal anti-inflammatory responses^{9–11}. However, comparatively little is known about the role of efferent vagus nerve signalling in modulating inflammation. Here, we describe a previously unrecognized, parasympathetic anti-inflammatory pathway by which the brain modulates systemic inflammatory responses to endotoxin. Acetylcholine, the principle vagal neurotransmitter, significantly attenuated the release of cytokines (tumour necrosis factor (TNF), interleukin (IL)-1 β , IL-6 and IL-18), but not the anti-inflammatory cytokine IL-10, in lipopolysaccharide-stimulated human macrophage cultures. Direct electrical stimulation of the peripheral vagus nerve *in vivo* during lethal endotoxaemia in rats inhibited TNF synthesis in liver, attenuated peak serum TNF amounts, and prevented the development of shock.

Vagus nerve signalling is a critical component of the afferent loop that modulates the adrenocorticotropin and fever responses to systemic endotoxaemia and cytokinaemia^{12–15}. Efferent vagus nerve signalling may facilitate lymphocyte release from thymus through a nicotinic acetylcholine receptor response¹⁶. Clinical studies indicate that nicotine administration can be effective for treating some cases of inflammatory bowel disease^{17,18}, and that proinflammatory cytokines are significantly decreased in the colonic mucosa of smokers with inflammatory bowel disease¹⁹. Accordingly, we reasoned that the cholinergic parasympathetic nervous system may modulate the systemic inflammatory response.

We established primary human macrophage cultures by incubating human peripheral blood mononuclear cells in the presence of macrophage colony stimulating factor (MCSF). Acetylcholine (ACh) inhibited TNF release dose-dependently in macrophage cultures conditioned by exposure to lipopolysaccharide (LPS) for 4 h (Fig. 1a). We observed a comparable inhibition of TNF release by ACh from macrophages exposed to LPS for 20 h (data not shown), indicating that ACh did not merely delay the onset of the TNF response. We also treated macrophage cultures with carbachol, a cholinergic agonist chemically distinct from ACh, and observed

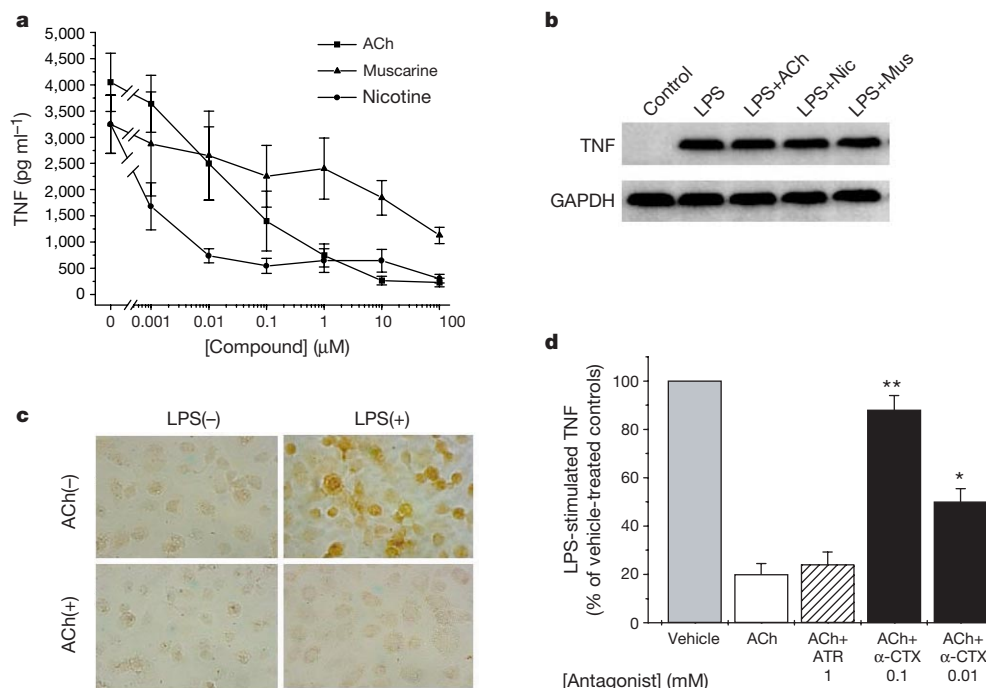


Figure 1 Cholinergic agonists inhibit LPS-induced TNF synthesis in human macrophage cultures through a post-transcriptional mechanism. **a**, Dose-dependent inhibition of TNF release by ACh (squares), muscarine (triangles) and nicotine (circles) after stimulation with LPS for 4 h. Data are mean $\pm$ s.e.m.; $n = 9$. **b**, ACh (100 μ M), muscarine (100 μ M, Mus) and nicotine (100 μ M, Nic) do not reduce amount of LPS-stimulated (2 h) TNF mRNA in macrophages. **c**, Immunostaining with anti-TNF antibodies reveals a substantial decrease

in LPS-induced (2 h) TNF immunoreactivity in ACh-treated (100 μ M) human macrophages. **d**, α -Conotoxin (α -CTX), but not atropine (ATR), restores the LPS-stimulated release of TNF in cultures treated with ACh (10 μ M). Neither atropine nor α -conotoxin altered TNF production in vehicle-treated cultures (not shown). Data shown are mean $\pm$ s.e.m. of three separate experiments. Asterisk, $P < 0.05$ versus ACh; double asterisk, $P < 0.005$ versus ACh.

inhibition of TNF release (data not shown). We investigated the molecular mechanism of TNF inhibition by measuring TNF messenger RNA amounts in an RNase protection assay. The amount of TNF mRNA in ACh-treated, LPS-stimulated macrophages did not decrease compared with vehicle-treated, LPS-stimulated macrophages, even when ACh was added in concentrations that inhibited TNF protein release (Fig. 1b). This indicates that ACh suppresses TNF through a post-transcriptional mechanism. To determine whether ACh inhibited TNF synthesis or release, we labelled cell-associated TNF in human macrophages using monoclonal anti-TNF antibodies. ACh significantly attenuated LPS-stimulated TNF immunoreactivity in macrophages (Fig. 1c), indicating that the inhibitory effect of ACh on human macrophage TNF production occurs through the post-transcriptional suppression of TNF protein synthesis.

Peripheral blood mononuclear cells express nicotinic and muscarinic ACh receptors^{20–22}. To define pharmacologically the type of macrophage cholinergic receptors involved in modulating the TNF response, we measured TNF in macrophage culture medium exposed to LPS for 4 h in the presence of either muscarinic or nicotinic agonists of ACh receptors (Fig. 1a). Nicotine significantly inhibited TNF release in a dose-dependent manner. The effective concentration of nicotine that inhibited 50% of the TNF response (EC_{50}) was estimated to be 8.3 ± 7.1 nM ($n = 9$). This compared favourably with the EC_{50} for ACh-mediated inhibition of TNF (ACh $EC_{50} = 20.2 \pm 8.7$ nM, $n = 9$). Muscarine also significantly inhibited TNF release, although it was much less effective than either ACh or nicotine (muscarine $EC_{50} = 42.4 \pm 18.6$ μ M, $n = 9$; $P < 0.01$ versus nicotine or ACh). As these results did not establish whether ACh inhibited TNF primarily through the activity of nicotinic or muscarinic ACh receptors, we added the specific muscarinic antagonist, atropine, to LPS-stimulated macrophage cultures co-treated with ACh (Fig. 1d). Addition of atropine, even in concen-

trations as high as 1 mM, failed to restore TNF release in ACh-treated macrophage cultures. We next addressed whether the nicotinic ACh receptor activity that mediated inhibition of TNF was sensitive to α -bungarotoxin. Addition of α -conotoxin to ACh-treated LPS-stimulated macrophage cultures significantly and dose-dependently reversed the inhibitory effect of ACh (Fig. 1d). These data indicate that the inhibitory effect of ACh on the LPS-induced TNF response in human macrophage cultures is mediated primarily by α -bungarotoxin-sensitive, nicotinic ACh receptors. The amount of ACh in mammalian tissues can reach the millimolar range²³, so it is possible that both the nicotinic and muscarinic macrophage ACh receptors participate in the inhibition of macrophage TNF synthesis *in vivo*.

To assess the specificity of the cytokine response to ACh, we measured the release of other macrophage-derived cytokines in LPS-stimulated, ACh-treated macrophage cultures. ACh dose-dependently inhibited the release of other LPS-inducible cytokines (IL-1 β , IL-6 and IL-18), but failed to prevent the constitutive release of the anti-inflammatory cytokine IL-10 (Fig. 2). Staining with tetrazolium salt MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide), and Trypan blue exclusion of macrophage cultures treated with LPS and ACh, indicated that specific LPS-inducible cytokine inhibition was not because of cytotoxicity (data not shown). The molecular mechanism of ACh inhibition of IL-1 β and IL-6 was investigated further by measuring gene-specific mRNA amounts with biotin-labelled capture oligonucleotide probes in a microplate assay (Quantikine mRNA, R&D Systems). Stimulation of human macrophage cultures with LPS for 2 h significantly increased the mRNA amounts of IL-1 β compared with vehicle-treated controls (vehicle-treated IL-1 β mRNA = 120 ± 54 attomole (μ mol) ml^{-1} versus LPS-stimulated IL-1 β mRNA = 1974 ± 179 (attomole) $attomole\ ml^{-1}$; $n = 3$; $P < 0.01$). Addition of ACh in concentrations that inhibited IL-1 β protein release (100 nM) did

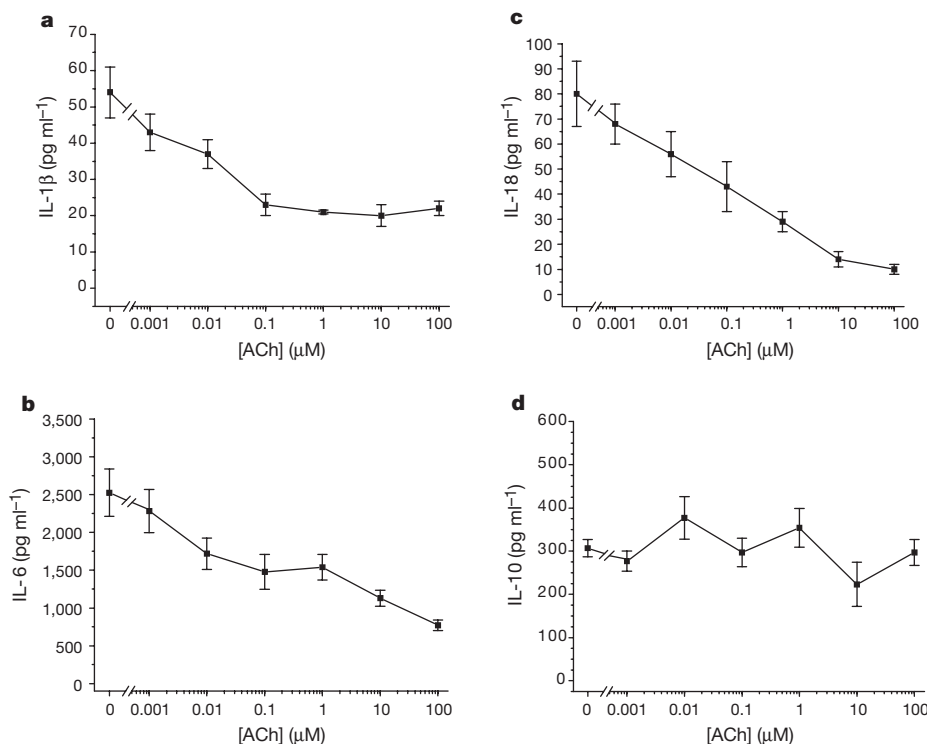


Figure 2 ACh specifically inhibits release of pro-inflammatory cytokines by human macrophages, but does not suppress release of the anti-inflammatory cytokine, IL-10. **a–d**, LPS-stimulated human macrophage cultures were incubated with ACh at the indicated concentrations in the presence of pyridostigmine bromide (1 mM) and LPS

(100 ng ml^{-1}). IL-1 β (**a**), IL-6 (**b**), IL-18 (**c**) and IL-10 (**d**) amounts were measured in conditioned media (20 h) using commercially available ELISA kits. Data are mean $\pm$ s.e.m. from four separate experiments.

not significantly alter macrophage IL-1 β mRNA amounts (ACh-treated LPS-stimulated IL-1 β mRNA = 2128 $\pm$ 65 amol ml $^{-1}$; n = 3). Similarly, LPS-stimulated IL-6 mRNA amounts in macrophages were not significantly altered by ACh concentrations that inhibited IL-6 protein (LPS-stimulated IL-6 mRNA = 1716 $\pm$ 157 amol ml $^{-1}$ versus ACh-treated LPS-stimulated IL-6 mRNA = 1872 $\pm$ 91 amol ml $^{-1}$; n = 3). These observations indicate that ACh post-transcriptionally inhibits the LPS-stimulated release of TNF, IL-1 β and IL-6 in macrophages.

To determine whether direct stimulation of efferent vagus nerve activity might suppress the systemic inflammatory response to endotoxin, we subjected adult male Lewis rats to bilateral cervical vagotomy, or a comparable sham surgical procedure in which the vagus nerves were isolated but not transected. Efferent vagus nerve activity was stimulated in vagotomized rats by application of constant voltage pulses to the distal end of the divided vagus nerve 10 min before and again 10 min after the administration of a lethal LPS dose (15 mg kg $^{-1}$, intravenous). Electrical stimulation of

the efferent vagus nerve significantly decreased the amounts of TNF in the serum. By contrast, vagotomy without electrical stimulation significantly increased peak serum TNF amounts compared with sham-operated controls (P < 0.05) (Fig. 3a). We next measured TNF in liver homogenates, because liver is a principal source of serum TNF during endotoxaemia^{24,25}. Electrical stimulation of the distal vagus nerve decreased hepatic LPS-stimulated TNF synthesis compared with sham-operated controls. Vagotomy without stimulation was associated with increased TNF synthesis in liver (Fig. 3b). These data directly implicate efferent vagus nerve signalling in the regulation of TNF production *in vivo*.

It was possible that electrical stimulation of the vagus nerve induced the release of humoral anti-inflammatory hormones or cytokines that inhibit TNF production. Measurement of corticosterone and IL-10 in sham-operated controls (Table 1) revealed that endotoxaemia was associated with increases in the amount of corticosterone and IL-10. In agreement with previous studies, vagotomy significantly reduced the amount of corticosterone, in part because it eliminated the afferent vagus nerve signals to the brain that are required for a subsequent activation of the hypothalamic–pituitary–adrenal signalling pathway^{12,13}. This decreased corticosteroid response probably contributed to the increased amounts of TNF observed in the serum and liver of vagotomized animals (Fig. 3) because corticosteroids normally downregulate TNF production^{10,11}. Direct electrical stimulation of the peripheral vagus nerve did not stimulate an increase in either the corticosteroid or the IL-10 responses. Thus, suppressed TNF synthesis in the serum and liver after vagus nerve stimulation could not be attributed to the activity of these humoral anti-inflammatory mediators.

Peripheral vagus nerve stimulation significantly attenuated the development of LPS-induced hypotension (shock) in rats exposed to lethal doses of endotoxin (Fig. 3c). This was expected, as TNF is a principle early mediator of acute endotoxin-induced shock^{1,26}. Vagotomy without stimulation significantly shortened the time to development of shock compared with sham-operated controls (time to 50% drop in mean arterial blood pressure: sham = 30 $\pm$ 3 min versus vagotomy = 15 $\pm$ 2 min; P < 0.05). This amplified development of shock following vagotomy corresponds to the decreased corticosteroid response and the increased TNF response. ACh is a vasodilator that mediates nitric oxide-dependent relaxation of resistance in blood vessels causing a decrease in blood pressure. We wished to exclude the possibility that stimulation of the efferent vagus might have mediated a paradoxical hypertensive response. Hypertension was not observed following vagus nerve stimulation of controls given saline instead of endotoxin (not shown), indicating that protection against endotoxic shock by vagus nerve stimulation is specific. These observations indicate that stimulation of efferent vagus nerve activity downregulates systemic TNF production and the development of shock during lethal endotoxaemia.

The present results show that differentiated human macrophage cultures are extremely sensitive to ACh and nicotine. Previous reports of cholinergic receptor activity in human peripheral blood mononuclear cells that were not differentiated into macrophages^{21,27,28}

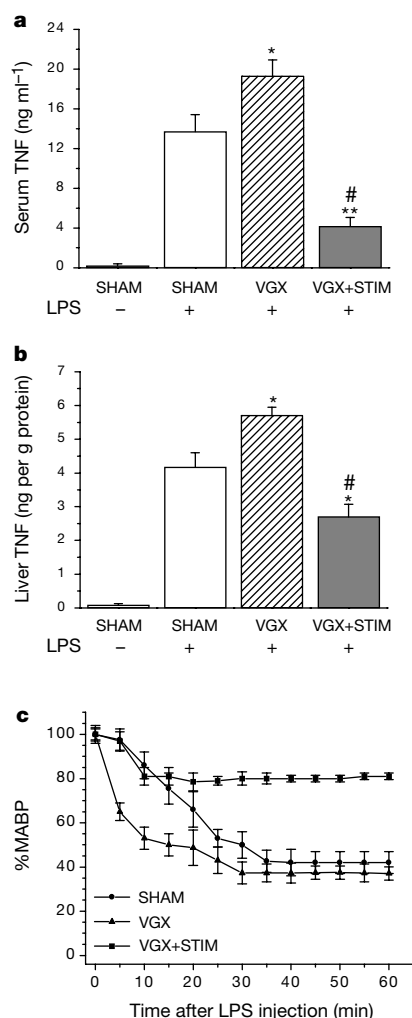


Figure 3 Vagus nerve stimulation attenuates the LPS-induced serum TNF response, hepatic TNF response, and development of endotoxic shock. Rats were subjected to either sham surgery (SHAM, white bars (a, b) or circles (c), n = 7), bilateral cervical vagotomy (VGX, striped bars (a, b) or triangles (c), n = 7) or vagotomy and electrical stimulation (VGX + STIM, shaded bars (a, b) or squares (c), n = 7). **a**, Serum TNF response. **b**, hepatic TNF response. **c**, Development of endotoxic shock. Mean arterial blood pressure data are normalized to MABP at time = 0. Sham-surgery, vagotomy and electrical stimulation with vagotomy did not significantly affect MABP in vehicle-treated controls (not shown). Data are mean $\pm$ s.e.m. Asterisk, P < 0.05; double asterisk, P < 0.005 versus SHAM + LPS; hash, P < 0.05 versus VGX + LPS.

Table 1 Effects of vagotomy and vagus nerve stimulation on serum IL-10 and corticosteroid amounts during lethal endotoxaemia

Group of animals	IL-10 (ng ml $^{-1}$)	Corticosterone (ng ml $^{-1}$)
SHAM + vehicle	n.d.	160 $\pm$ 20
SHAM + LPS	8 $\pm$ 0.3	850 $\pm$ 50
Vagotomy + LPS	9 $\pm$ 0.4	570 $\pm$ 34*
Vagotomy + LPS + Stimulation	9 $\pm$ 0.5	560 $\pm$ 43*

Animals were subjected to either sham surgery (SHAM), vagotomy (VGX), or electrical stimulation with vagotomy (VGX+STIM) 30 min before systemic administration of LPS (15 mg kg $^{-1}$). Blood samples were collected 1 h after administration of LPS or vehicle. Serum corticosterone was measured by radioimmunoassay (ICN Biomedicals, Costa Mesa, CA) and IL-10 was determined by ELISA (BioSource International, Camarillo, CA). All assays were performed in triplicate. Data shown are mean $\pm$ s.e.m., n = 7 animals per group. Asterisk, P < 0.05 versus SHAM + LPS; n.d., not detectable.

implied that maximal cholinergic responses required micromolar concentrations of cholinergic agonists. Our own studies of undifferentiated human peripheral blood mononuclear cells confirmed that significantly higher concentrations of ACh are required to suppress cytokine synthesis (ACh EC₅₀ for inhibiting TNF in peripheral blood mononuclear cells = 0.8 ± 0.2 mM, $n = 3$). The pharmacological results now implicate an α -bungarotoxin-sensitive, nicotinic ACh receptor activity that can modulate the macrophage cytokine response. This type of cholinergic receptor activity is similar to that previously described in peripheral blood mononuclear cells²¹, except that macrophages are significantly more sensitive to cholinergic agonists than peripheral blood mononuclear cells.

The neural-immune interaction described here—which we call the ‘cholinergic anti-inflammatory pathway’—can directly modulate the systemic response to pathogenic invasion. The observation that parasympathetic nervous system activity influences circulating TNF amounts and the shock response to endotoxaemia has widespread implications, because it is a previously unrecognized, direct and rapid endogenous mechanism that can suppress the lethal effects of biological toxins. The cholinergic anti-inflammatory pathway has much shorter response times than the humoral anti-inflammatory pathways. Moreover, activation of parasympathetic efferents during systemic stress, or the ‘fight or flight’ response, confers an additional protective advantage to the host by restraining the magnitude of a potentially lethal peripheral immune response. □

Methods

Human macrophage cultures

Buffy coats were collected from the blood of healthy individual donors to the Long Island Blood Bank Services (Melville, New York). Primary blood mononuclear cells were isolated by density-gradient centrifugation through Ficoll/Hypaque (Pharmacia), suspended (8×10^6 cells ml⁻¹) in RPMI 1640 medium with 10% heat inactivated human serum (Gemini Bio-Products), and seeded in flasks. After incubation for 2 h at 37 °C, adherent cells were detached with 10 mM EDTA, resuspended (10^6 cells ml⁻¹) in medium supplemented with human MCSF (Sigma; 2 ng ml⁻¹), and seeded onto 24-well tissue culture plates (10^6 cells per well). Cells were allowed to differentiate for seven days in the presence of MCSF.

On day seven, fresh medium without MCSF was added, and experiments performed as indicated. Human macrophage cultures were exposed to ACh chloride (Sigma), muscarine or nicotine where indicated. All experiments using ACh included the ACh esterase inhibitor pyridostigmine bromide (1 mM, Sigma). Five minutes after the addition of the cholinergic agents, LPS was added to the cultures at a final concentration of 100 ng ml⁻¹. Supernatants were collected 4 or 20 h after the addition of LPS and prepared for cytokine ELISA in accordance with the manufacturers’ instructions. ELISA kits were obtained from R&D Systems (TNF, IL-1 β , IL-6 and IL-10) and Medical & Biological Laboratories (IL-18). For experiments using atropine (Sigma) or α -conotoxin (Oncogene), the cholinergic antagonists were added to the macrophage cultures 5 min before ACh (10 μ M) and LPS.

RNase protection assay

Macrophages were incubated with 100 μ M ACh, muscarine (Mus) or nicotine (Nic) for 5 min followed by 2 h exposure to LPS. Control wells had medium alone. The RNase protection assay was conducted using a kit obtained from PharMingen. The anti-sense RNA probe set (hck-3) was labelled with [α -³²P]UTP (800 Ci mmol⁻¹, Amersham) using T7 RNA polymerase. Molecular weight markers were prepared by using pBR-322 plasmid DNA digested with *MspI* (New England Bio Labs) and end-labelled using [α -³²P]dCTP (800 Ci mmol⁻¹, Amersham) with Klenow enzyme (Stratagene). Expression of the GAPDH gene product was measured to control for mRNA loading. Total RNA was isolated from cultured cells by using TRIzol reagent (GIBCO BRL) following the manufacturer’s instructions.

TNF immunohistochemistry

Human macrophages were differentiated as described above, and grown on glass chamber slides. Cells were exposed to ACh (100 μ M) in the presence of pyridostigmine bromide (1 mM), 5 min before LPS treatment. Two hours later the cells were fixed in buffered 10% formalin and subjected to immunocytochemical analysis. Slides were incubated in a blocking solution (1% BSA, 5% normal goat serum, 0.3% Triton X-100 in PBS) for 1 h at room temperature and then incubated for 24 h at 4 °C with a primary mouse anti-human TNF monoclonal antibody (Genzyme) diluted 1:100 in PBS containing 0.3% Triton X-100, 0.1% BSA, and 3% normal goat serum. Washed sections were incubated for 2 h with secondary biotinylated anti-mouse IgG (1:200, Vector Laboratories). The reaction product was visualized with 0.003% hydrogen peroxide and 0.05% 3,3’-diaminobenzidine

tetrahydrochloride as a chromogen. Negative controls were incubated in the absence of primary antibodies (not shown). Slides were analysed on a light microscope (Olympus BX60) using a MetaMorph Imaging System (Universal Imaging).

Animal model of endotoxic shock

Adult male Lewis rats (280–300 g, Charles River) were housed at 22 °C on a 12 h light/dark cycle. All animal experiments were performed in accordance with the National Institute of Health Guidelines under the protocols approved by the Institutional Animal Care and Use Committee of North Shore University Hospital/ New York University School of Medicine. Rats were anaesthetized with urethane (1 g kg⁻¹, intraperitoneally), and the trachea, the common carotid artery, and the jugular vein were cannulated with polyethylene tubing (Clay Adams). The catheter implanted into the right common carotid artery was connected to a blood pressure transducer and an Acquisition System (MP100, BIOPAC Systems) for continuous registration of mean arterial blood pressure (MABP). Animals were subjected to bilateral cervical vagotomy (VGX, $n = 7$) alone or with electrical stimulation (VGX+STIM, $n = 7$) or sham surgery (SHAM, $n = 7$). In vagotomized animals, following a ventral cervical midline incision, both vagus trunks were exposed, ligated with a 4-0 silk suture, and divided. In sham-operated animals both vagal trunks were exposed and isolated from the surrounding tissue but not transected. Electrical stimulation of the vagus nerve was performed in vagotomized animals. In these groups, the distal end of a vagus nerve trunk was placed across bipolar platinum electrodes (Plastics One) connected to a stimulation module (STM100A, Harvard Apparatus) and controlled by an Acquisition System. Constant voltage stimuli (5 V, 2 ms, 1 Hz) were applied to the nerve for 20 min (10 min before LPS administration and 10 min after). LPS (*Escherichia coli* 0111:B4; Sigma; 10 mg ml⁻¹ in saline) was sonicated for 30 min, and administered at a lethal dose (15 mg kg⁻¹, intravenously). Blood was collected from the right carotid artery 1 h after LPS administration. Serum TNF amounts were quantified by the L929 bioactivity assay. To determine liver TNF amounts, animals were killed and livers rapidly excised, rinsed of blood, homogenized by polytron (Brinkman) in homogenization buffer (PBS, containing 0.05 % sodium azide, 0.5% Triton X-100 and a protease inhibitor cocktail; pH 7.2; 4 °C), and then sonicated for 10 min. Homogenates were centrifuged at 12,000g for 10 min, and TNF amounts in supernatants determined by ELISA (Biosource International). Liver TNF content was normalized to the amount of protein in the sample measured by the Bio-Rad protein assay.

Statistical analysis

Data shown are mean $\pm$ s.e.m. Statistical analyses between groups were performed using the two-tailed Student’s *t*-test; *P* values < 0.05 were considered significant.

Received 12 January; accepted 3 March 2000.

- Tracey, K. J. *et al.* Shock and tissue injury induced by recombinant human cachectin. *Science* **234**, 470–474 (1986).
- Dinarello, C. A. The interleukin-1 family: 10 years of discovery. *FASEB J.* **8**, 1314–1325 (1994).
- Wang, H. *et al.* HMGI as a late mediator of endotoxin lethality in mice. *Science* **285**, 248–251 (1999).
- Nathan, C. F. Secretory products of macrophages. *J. Clin. Invest.* **79**, 319–326 (1987).
- Besedovsky, H., Rey, D. A., Sorkin, E. & Dinarello, C. A. Immunoregulatory feedback between interleukin-1 and glucocorticoid hormones. *Science* **233**, 652–654 (1986).
- Woiciechowsky, C. *et al.* Sympathetic activation triggers systemic interleukin-10 release in immunodepression induced by brain injury. *Nature Med.* **4**, 808–813 (1998).
- Hu, X. X., Goldmuntz, E. A. & Brosnan, C. F. The effect of norepinephrine on endotoxin-mediated macrophage activation. *J. Neuroimmunol.* **31**, 35–42 (1991).
- Lipton, J. M. & Catania, A. Anti-inflammatory actions of the neuroimmunomodulator alpha-MSH. *Immunol. Today* **18**, 140–145 (1997).
- Watkins, L. R. & Maier, S. F. Implications of immune-to-brain communication for sickness and pain. *Proc. Natl Acad. Sci. USA* **96**, 7710–7713 (1999).
- Sternberg, E. M. Neural-immune interactions in health and disease. *J. Clin. Invest.* **100**, 2641–2647 (1997).
- Scheinman, R. I., Cogswell, P. C., Lofquist, A. K. & Baldwin, A. S. Jr Role of transcriptional activation of IkappaB alpha in mediation of immunosuppression by glucocorticoids. *Science* **270**, 283–286 (1995).
- Gaykema, R. P., Dijkstra, I. & Tilders, F. J. Subdiaphragmatic vagotomy suppresses endotoxin-induced activation of hypothalamic corticotropin-releasing hormone neurons and ACTH secretion. *Endocrinology* **136**, 4717–4720 (1995).
- Fleshner, M. *et al.* Thermogenic and corticosterone responses to intravenous cytokines (IL-1beta and TNF-alpha) are attenuated by subdiaphragmatic vagotomy. *J. Neuroimmunol.* **86**, 134–141 (1998).
- Watkins, L. R. *et al.* Blockade of interleukin-1 induced hyperthermia by subdiaphragmatic vagotomy: evidence for vagal mediation of immune-brain communication. *Neurosci. Lett.* **183**, 27–31 (1995).
- Romanovsky, A. A., Simons, C. T., Szekeley, M. & Kulchitsky, V. A. The vagus nerve in the thermoregulatory response to systemic inflammation. *Am. J. Physiol.* **273**, R407–R413 (1997).
- Antonica, A., Magni, F., Mearini, L. & Paolucci, N. Vagal control of lymphocyte release from rat thymus. *J. Auton. Nerv. Syst.* **48**, 187–197 (1994).
- Guslandi, M. Nicotine treatment for ulcerative colitis. *Br. J. Clin. Pharmacol.* **48**, 481–484 (1999).
- Sandborn, W. J. *et al.* Transdermal nicotine for mildly to moderately active ulcerative colitis. A randomized, double-blind, placebo-controlled trial. *Ann. Intern. Med.* **126**, 364–371 (1997).
- Sher, M. E. *et al.* The influence of cigarette smoking on cytokine levels in patients with inflammatory bowel disease. *Inflamm. Bowel Dis.* **5**, 73–78 (1999).
- Sato, E., Koyama S., Okubo, Y., Kubo, K. & Sekiguchi, M. Acetylcholine stimulates alveolar macrophages to release inflammatory cell chemotactic activity. *Am. J. Physiol.* **274**, L970–L979 (1998).
- Whaley, K., Lapping, D. & Barkas, T. C2 synthesis by human monocytes is modulated by a nicotinic cholinergic receptor. *Nature* **293**, 580–583 (1981).
- Sato, K. Z. *et al.* Diversity of mRNA expression for muscarinic acetylcholine receptor subtypes and neuronal nicotinic acetylcholine receptor subunits in human mononuclear leukocytes and leukemic cell lines. *Neurosci. Lett.* **266**, 17–20 (1999).

23. Wathey, J. C., Nass, M. M. & Lester, H. A. Numerical reconstruction of the quantal event at nicotinic synapses. *Biophys. J.* **27**, 145–164 (1979).
24. Kumins, N. H., Hunt, J., Gamelli, R. L. & Filkins, J. P. Partial hepatectomy reduces the endotoxin-induced peak circulating level of tumor necrosis factor in rats. *Shock* **5**, 385–388 (1996).
25. Gregory, S. H. & Wing, E. J. Neutrophil-Kupffer-cell interaction in host defenses to systemic infections. *Immunol. Today* **19**, 507–510 (1998).
26. Tracey, K. J. *et al.* Anti-cachectin/TNF monoclonal antibodies prevent septic shock during lethal bacteraemia. *Nature* **330**, 662–664 (1987).
27. Madretsma, G. S. *et al.* Nicotine inhibits the in vitro production of interleukin 2 and tumour necrosis factor- α by human mononuclear cells. *Immunopharmacology* **35**, 47–51 (1996).
28. van Dijk, A. P. *et al.* Transdermal nicotine inhibits interleukin 2 synthesis by mononuclear cells derived from healthy volunteers. *Eur. J. Clin. Invest.* **28**, 664–671 (1998).

Acknowledgements

We thank D. Nardi for help in ELISA assays and D. Prieto for administrative assistance.

Correspondence and requests for materials should be addressed to L.V.B.

(e-mail: lborovikova@mindspring.com) or K.J.T. (e-mail: kjtracey@sprynet.com).

Targeted destabilization of HY5 during light-regulated development of *Arabidopsis*

Mark T. Osterlund, Christian S. Hardtke, Ning Wei & Xing Wang Deng

Department of Molecular, Cellular and Developmental Biology, Yale University, PO Box 208104, 165 Prospect Street, OML 301, New Haven, Connecticut 06520-8104, USA

***Arabidopsis* seedlings display contrasting developmental patterns depending on the ambient light. Seedlings grown in the light develop photomorphogenically, characterized by short hypocotyls and expanded green cotyledons. In contrast, seedlings grown in darkness become etiolated, with elongated hypocotyls and closed cotyledons on an apical hook. Light signals, perceived**

by multiple photoreceptors and transduced to downstream regulators, dictate the extent of photomorphogenic development in a quantitative manner. Two key downstream components, COP1 and HY5, act antagonistically in regulating seedling development¹. HY5 is a bZIP transcription factor that binds directly to the promoters of light-inducible genes, promoting their expression and photomorphogenic development^{2,3}. COP1 is a RING-finger protein with WD-40 repeats whose nuclear abundance is negatively regulated by light^{4,5}. COP1 interacts directly with HY5 in the nucleus to regulate its activity negatively¹. Here we show that the abundance of HY5 is directly correlated with the extent of photomorphogenic development, and that the COP1–HY5 interaction may specifically target HY5 for proteasome-mediated degradation in the nucleus.

To characterize the nuclear HY5 protein^{2,3} biochemically, we produced rabbit polyclonal antibodies (anti-HY5) against recombinant HY5 protein. Western blot analyses of total protein extracts from light-grown wild-type seedlings revealed that anti-HY5 recognizes a protein band at an apparent relative molecular mass of 30,000 (M_r 30K) (Fig. 1a). The apparently large size of HY5 relative to its predicted size of 18K (ref. 2) is largely due to abnormal migration of the HY5 protein in our polyacrylamide gel system (data not shown). The presence of a 30K band in wild-type seedlings but not in the three *hy5* mutant alleles examined indicates that it is the endogenous *Arabidopsis* HY5.

Quantitative western blot analyses indicated that HY5 is 15–20 times more abundant in seedlings grown in white light than in those grown in the dark (Fig. 1b). This difference is observed in green tissues (cotyledon and hypocotyl) as well as in the non-photosynthetic roots (data not shown). We examined HY5 levels during light–dark transitions. Wild-type seedlings were grown in continuous light or darkness for four days and then transferred to the opposite light condition for 5, 10, 15 or 20 h. A western blot of the protein extracts from these seedlings revealed that although changes in the abundance of HY5 occur within 5 h, a full day of light or darkness is required for HY5 to reach its maximum or minimum, respectively (Fig. 1c).

Consistent with a previous report², HY5 messenger RNA levels

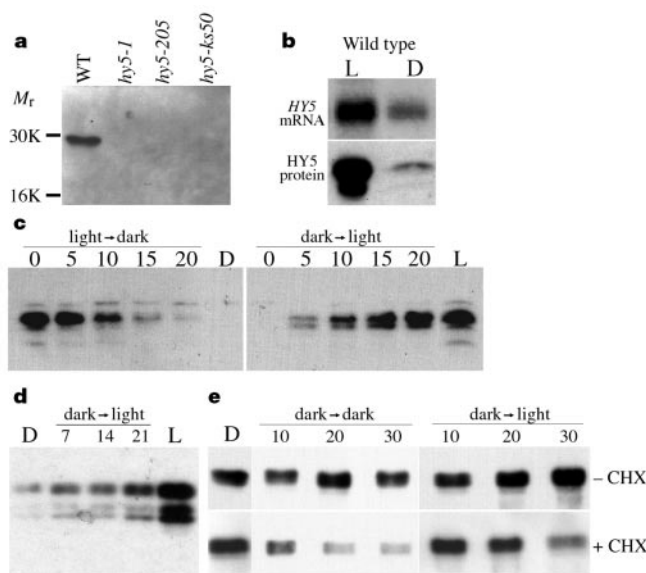


Figure 1 HY5 expression is regulated by light at both the mRNA and protein levels. **a**, A western blot of wild-type (WT), *hy5-1*, *hy5-205* and *hy5-ks50* seedling extracts using anti-HY5. **b**, A *HY5* northern blot and an anti-HY5 western blot of light- and dark-grown seedlings (L, light; D, dark). **c**, Anti-HY5 western blots from seedlings grown in continuous light (L) or darkness (D) for 4 d and then transferred to the opposite light condition for 5, 10, 15 or 20 h. **d**, An anti-HY5 western blot of 35SHY5 seedlings grown in continuous

darkness for 4 d and then transferred to light for 0 (D), 7, 14 or 21 h. Seedlings grown in continuous light were included as a positive control (L). **e**, Anti-HY5 western blots of 35SHY5 seedlings treated with cycloheximide (+CHX) or buffer only (–CHX). Seedlings were grown in continuous darkness for four days before treatment and then transferred to light or darkness. Protein extracts were made 0, 10, 20 and 30 h after the addition of cycloheximide.

showed only a two-to-threefold difference between light- and dark-grown seedlings (Fig. 1b). This is considerably less than the 15–20-fold difference seen in HY5 protein levels. Therefore the mRNA levels could contribute to the differential accumulation of the HY5 protein, but translational or post-translational regulation may be primarily responsible for regulating the abundance of HY5. This conclusion is supported by the observation that HY5 exhibits a similar light-induced accumulation in HY5-overexpressing (35SHY5) transgenic seedlings (Fig. 1d). This is particularly convincing considering the greater than 50-fold excess in the total HY5 mRNA and its lack of light regulation in the transgenic lines (data not shown).

We used cycloheximide, a cytoplasmic protein-synthesis inhibitor, to determine whether protein degradation or protein synthesis is the key regulatory step for the differential abundance of HY5 in the light and dark. Four-day-old 35SHY5 seedlings grown in darkness were transferred to liquid cultures containing cycloheximide. These cultures were then placed in white light or darkness for variable lengths of time, after which the abundance of HY5 was determined. For reasons not yet clear, the HY5 protein increased more slowly and to a lesser extent in seedlings grown in liquid culture (Fig. 1e, top left) than in seedlings grown on plates (Fig. 1d). Although the amount of HY5 decreases within 30 h in both light and darkness in the presence of cycloheximide, the rate of decrease is significantly greater in darkness (Fig. 1e, bottom). As 100 μ M cycloheximide effectively blocks new protein synthesis (data not shown), the reduced levels of HY5 must be due to the degradation of pre-existing protein. Therefore, our result indicates that HY5 is degraded at a greater rate in darkness than in light.

To reveal which cellular mechanism is responsible for HY5 degradation, we developed a cell-free assay based on a reported degradation system⁶. In this assay, HY5 is completely degraded within one hour (Fig. 2a). This degradation is greatly enhanced by the presence of exogenous ATP (data not shown) and can be prevented by four distinct proteasome-specific inhibitors (Fig. 2b). However, general protease inhibitors (leupeptin, PMSF) had little or no effect on the degradation of HY5 in this assay. As the control proteins⁷ showed little or no degradation in the extract during the assay, the degradation of HY5 was specific. Thus, we conclude that the proteasome is largely responsible for the degradation of HY5.

Eleven pleiotropic *COP/DET/FUS* loci are required to mediate repression of photomorphogenic development^{8,9}. Of these, at least six *COP/DET/FUS* proteins are components of the COP9 signalosome, a multi-subunit complex that is similar to the lid subcomplex

of the 19S proteasome regulatory particle^{10–13}. This structural similarity, together with the reported interaction between subunits of the COP9 signalosome and the proteasome⁷, may indicate that the COP9 signalosome has a direct or regulatory role in proteasome-mediated degradation. Therefore we examined the light regulation of HY5 in representative mutants of all available pleiotropic *COP/DET/FUS* loci. Although there was still a two-to-threefold difference in HY5 mRNA levels between light- and dark-grown mutant seedlings (Fig. 3 and ref. 2), HY5 protein accumulated to similar levels in both light and darkness in all mutants examined (Fig. 3). Furthermore, *in vitro* assays using extract from the weak *cop1-6* mutant seedlings showed a greatly reduced rate of proteasome-mediated degradation of HY5 protein (data not shown). Therefore, all defined pleiotropic *COP/DET/FUS* proteins are essential for the light-regulated instability of HY5 that is mediated by the proteasome.

The accumulation of or reduction in HY5 during light–dark transitions occurred over the same timescale as the redistribution of COP1 between the cytoplasm and the nucleus¹⁴. As nuclear COP1 interacts with HY5 and negatively regulates its activity¹⁵, COP1 may directly regulate HY5 stability. As COP1 activity and nuclear abundance are controlled by a wide spectrum of light through several photoreceptors¹⁶, we first examined whether the same set of photoreceptors is also involved in regulating the abundance of HY5. To determine their effects on the accumulation of HY5, photoreceptor mutant and overexpression lines were grown in various wavelengths of light and the abundance of HY5 was determined (Fig. 4). As expected, phytochrome B¹⁷, phytochrome A¹⁸ and the cryptochromes (CRY1 and CRY2)^{19,20} are the primary photoreceptors responsible for mediating the accumulation of HY5 under specific light. CRY1 and CRY2 are largely redundant and are responsible for the accumulation of HY5 in blue light (Fig. 4c). In red light, phyB is primarily responsible for the accumulation of HY5, although phyA has a minor role (Fig. 4a). Interestingly, phyA and the two cryptochromes all affect the abundance of HY5 in far-red light (Fig. 4b). In all cases, the abundance of HY5 directly correlates with the degree of photomorphogenic development. There is also a strong correlation between the effects of each photoreceptor on the nuclear exclusion of COP1 (ref. 16) and HY5 abundance.

To provide further evidence that nuclear COP1 is involved in regulating HY5 abundance, we determined the levels of HY5 in wild-type seedlings grown in various intensities of continuous white light. Figure 5a shows that in high-intensity light, seedlings display typical photomorphogenic phenotypes and HY5 accumulates to high levels. As the intensity decreases, HY5 also decreases and seedlings become more etiolated with longer hypocotyls and smaller cotyledons. Increasing light intensities result in a proportional decrease in the nuclear abundance of COP1 (refs 5, 14). Thus the

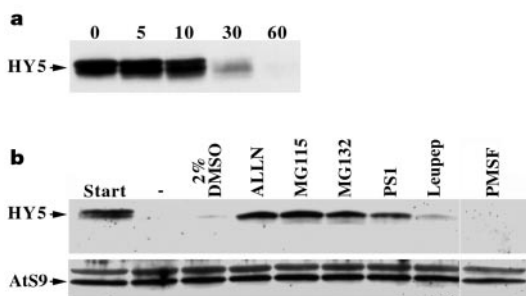


Figure 2 HY5 degradation is mediated by the proteasome pathway. **a**, Cell-free degradation of HY5 is completed within 1 h in a buffer system supporting the proteasome pathway^{6,30}. **b**, Cell-free degradation of HY5 in the presence of different inhibitors. Extracts were treated with nothing (–), 2% DMSO (solvent for the inhibitors), 40 μ M of one of the proteasome inhibitors ALLN, MG115, MG132 or PSI, 40 μ M leupeptin or 4 mM PMSF. After 2 h of incubation at room temperature, the HY5 levels were examined by western blotting. An unrelated control (AtS9) and a crossreacting band (asterisk) that are unaffected by proteasome degradation in the same extracts are shown as loading controls.

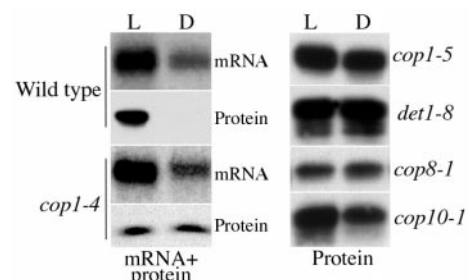


Figure 3 The pleiotropic *COP/DET/FUS* genes are essential for mediating light control of HY5 protein degradation but not its mRNA accumulation. Representative HY5 northern and western blots of wild-type, *cop1-4*, *cop1-5*, *det1-8*, *cop8-1* and *cop10-1* mutants grown in continuous light and darkness. All pleiotropic *cop/det/fus* mutants from the eleven available loci were examined and failed to reduce the abundance of HY5 in darkness. The blots were exposed for variable times for optimal results.

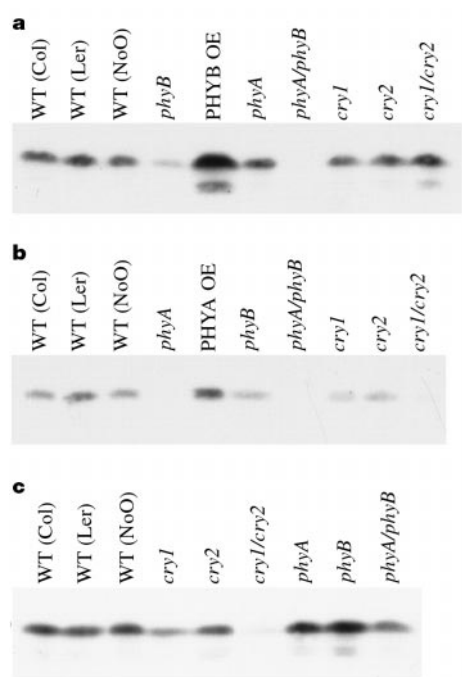


Figure 4 Several photoreceptors mediate the accumulation of HY5 under distinct wavelengths of light. **a–c**, Anti-HY5 western blots of seedlings grown in continuous red light (**a**), continuous far-red light (**b**) and continuous blue light (**c**). The seedlings include

wild type (three ecotypes), *cry1*, *cry2*, a *cry1/cry2* double mutant, *phyA*, *phyB*, a *phyA/phyB* double mutant, a PHYB overexpression line (PHYBOE, red light only) and a PHYA overexpression line (PHYAOE, far-red light only).

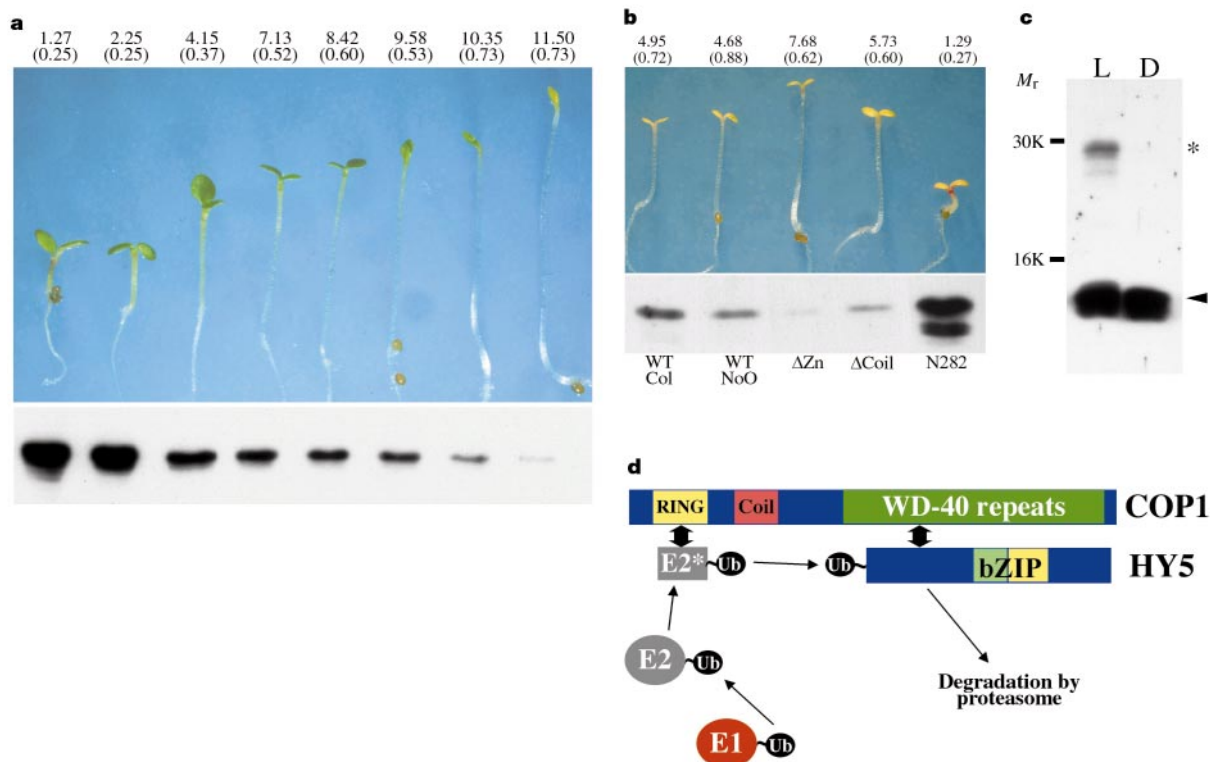


Figure 5 The light regulation of HY5 is dependent on the nuclear abundance of COP1 and its ability to interact with COP1. **a**, An anti-HY5 western blot of wild-type seedlings grown in varying light intensities (left to right: 156, 70, 30, 10, 5, 2.5, 0.6 and 0.3 $\mu\text{mol m}^{-2} \text{s}^{-1}$). Representative seedlings are shown above the western blot with the average hypocotyl lengths (s.d. in parentheses). **b**, An anti-HY5 western blot of wild-type (Columbia and No-O ecotypes), ΔZn , ΔCoil and N282 seedlings grown in continuous far-red light. Representative seedlings are shown above the western blot with the average hypocotyl

lengths (s.d. in parentheses). **c**, An anti-HY5 western blot of light- and dark-grown seedlings expressing truncated HY5 (ΔN77). Asterisk, wild-type HY5; arrowhead, ΔN77 . **d**, Proposed role of COP1 in targeting HY5 degradation. Here, COP1 acts as an ubiquitin-protein ligase (E3) that recruits an E2 and HY5 through distinct interacting domains (RING-finger motif and WD-40 repeat domain), mediating the ubiquitination of HY5 and its subsequent degradation by the proteasome.

nuclear abundance of COP1 is inversely correlated with the cellular abundance of HY5, which is exclusively localized within the nucleus^{2,3}.

To test whether the COP1–HY5 interaction is required for the observed degradation of HY5, we examined several transgenic lines expressing truncated versions of COP1 (Fig. 5b). HY5 interacts with COP1 through the WD-40 repeat domain^{1,21}. The overexpression of a truncated form of COP1 that lacks the WD-40 repeat domain (N282) results in a hyper-photomorphogenic phenotype^{21,22}. In contrast, overexpression of domain-deletion constructs of COP1 that lack the RING-finger domain (Δ Zn) or coiled-coil domain (Δ Coil) results in elongated hypocotyls²¹. The effects of the transgenes were most evident when seedlings were grown in far-red light (Fig. 5b, top). The abundance of HY5 in these transgenic lines was strictly correlated with the observed extent of photomorphogenic development. When grown in the same far-red light condition, both Δ Zn and Δ Coil had reduced levels of HY5 protein relative to wild type, whereas N282 showed a marked increase in HY5 (Fig. 5b, bottom).

We also examined the light-regulated accumulation of a truncated form of HY5 (Δ N77). Δ N77 lacks the COP1-interaction domain but still promotes photomorphogenic development, causing a hyper-photomorphogenic phenotype¹. As shown in Fig. 5c, the Δ N77 protein is not light-regulated and accumulates to similar levels in both light and darkness. In the same seedlings, the endogenous wild-type HY5 protein exhibits consistent light-regulated accumulation. This indicates that the amino-terminal portion of HY5, which contains the COP1-interacting domain, may be responsible for targeting HY5 for degradation. Those results support the conclusion that the interaction between COP1 and HY5 is necessary and essential for the degradation of HY5.

Our results show that the cellular abundance of HY5 is light regulated and is directly correlated with the degree of photomorphogenic development of the seedling. This regulation is primarily controlled at the level of protein degradation, probably mediated by the proteasome pathway. The data indicate that *Arabidopsis* COP1, a light-inactivatable repressor of photomorphogenesis, interacts directly with HY5 within the nucleus, resulting in degradation of HY5. Light may inactivate COP1, reducing its nuclear abundance and allowing the accumulation of HY5. Therefore, COP1 may directly target HY5 for proteasome-mediated degradation.

There is evidence that many RING-finger proteins can act as ubiquitin–protein ligases (E3) to target proteins for degradation by recruiting ubiquitin–conjugating enzymes (E2) and transferring the polyubiquitin from E2 to the targeted proteins^{23,24}. COP1 has both a RING-finger motif and a WD-40 repeat domain that can interact with several protein targets including HY5 (refs 15, 21). Thus it is not unreasonable to consider COP1, either alone or with other proteins, acting in a similar manner to an E3 ubiquitin–protein ligase. Specifically, nuclear COP1 may recruit an E2 ubiquitin–conjugating enzyme, targeting HY5 and other substrates for ubiquitination and subsequent degradation by the proteasome (Fig. 5d). The fact that all other pleiotropic COP/DET/FUS proteins are required for the light-regulated degradation of HY5 and that many of these proteins are part of the COP9 signalosome²⁵ indicates that all COP/DET/FUS proteins may have either a direct or a regulatory role in this proteasome-mediated degradation. □

Methods

Plant material and light conditions

The wild type used was of the Wassilewskija (WS) ecotype, unless otherwise indicated. The *cop/det/fus* mutants used have been described²⁵. The photoreceptor single and double mutants used were *phyA-201*, *phyB-5*, *phyA-201/phyB-5*, *cry1-304*, *cry2-1* and *cry1-304/cry2-1* (refs 17–19, 24, 26, 27). The PHYA overexpression lines (PHYAOE) and the PHYB overexpression lines (PHYBOE) were as described^{28,29}.

For all light shift experiments, seedlings were grown in continuous light or darkness for at least 4 d. Seedlings were then transferred to the opposite light condition for the designated length of time, such that all seedlings were the same age when protein was

extracted. Unless otherwise stated, the white light intensity used was $156 \mu\text{mol m}^{-2} \text{s}^{-1}$. The colour light growth chambers (Percival Scientific E-30LED2/3) have intensities of $33.5 \mu\text{mol m}^{-2} \text{s}^{-1}$ for blue light, $111.5 \mu\text{mol m}^{-2} \text{s}^{-1}$ for red light and $191.7 \mu\text{mol m}^{-2} \text{s}^{-1}$ for far-red light.

HY5 antibody and western analysis

We generated polyclonal HY5 antibodies by cloning the HY5 complementary DNA into the pet28a expression vector (Novagen). His₆-tagged HY5 was injected into rabbits at three-week intervals. HY5 antibody was purified from rabbit serum using GST–HY5 coupled to an *N*-hydroxy-succinimide-activated affinity column (Pharmacia Biotech).

All seedlings grown in the designated light conditions were 4–5 d old when protein was extracted. In most cases, 50 seedlings were ground in 50 μl of grinding buffer (400 mM sucrose, 50 mM Tris pH 7.5, 2.5 mM EDTA and 10 mM PMSF) and centrifuged, and equal amounts of extract were added to equal gel-loading buffer. Equal volumes of extract were loaded on the gels and western blots were performed.

Northern analysis

We isolated RNA from seedlings using the RNeasy kit (Qiagen). Northern blots were prepared using RNA from equal numbers of seedlings probed with radioactively labelled HY5 DNA fragments according to standard procedures.

Cycloheximide experiments

We grew 35SHY5 seedlings on germination media²² plates for 4 d in continuous darkness. After 4 d, seedlings were transferred to liquid culture (with the same germination medium lacking agar) where cycloheximide (100 μM) or buffer only (control) was added. The cultures were placed in the light or in the darkness and seedlings were removed 0, 10, 20 and 30 h after cycloheximide was added. Protein extracts were made by grinding seedlings in 100 μl of grinding buffer. The protein concentration was determined by Bradford assay and equal amounts of protein were loaded on an SDS–polyacrylamide gel.

Cell-free degradation assay

Five-day-old light-grown seedlings were ground in liquid nitrogen and resuspended in a buffer (25 mM Tris pH 7.5, 10 mM MgCl₂, 5 mM DTT, 10 mM NaCl and 10 mM ATP) modified according to ref. 30. Cell debris was pelleted by centrifugation and equal amounts of extract were transferred to individual tubes. For inhibitor studies, we incubated the extract aliquots at room temperature for 2 h in the presence of the respective supplements (Calbiochem), and reactions were stopped by adding an equal volume of 2 $\times$ protein gel-loading buffer. Equal amounts of sample were then analysed by western blot.

Received 1 November 1999; accepted 14 March 2000.

- Ang, L. H. *et al.* Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development. *Mol. Cell* **1**, 213–222 (1998).
- Oyama, T. *et al.* The *Arabidopsis* HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes Dev.* **11**, 2983–2995 (1997).
- Chattopadhyay, S. *et al.* *Arabidopsis* bZIP protein HY5 directly interacts with light-responsive promoters in mediating light control of gene expression. *Plant Cell* **10**, 673–683 (1998).
- Deng, X. W. *et al.* COP1, an *Arabidopsis* regulatory gene, encodes a protein with both a Zinc-binding motif and a G β homologous domain. *Cell* **71**, 791–801 (1992).
- von Arnim, A. G. & Deng, X. W. Light inactivation of *Arabidopsis* photomorphogenic repressor COP1 involves a cell-specific regulation of its nucleocytoplasmic partitioning. *Cell* **79**, 1035–1045 (1994).
- Naidoo, N., Song, W., Hunter-Ensor, M. & Sehgal, A. A role for the proteasome in the light response of the timeless clock protein. *Science* **285**, 1737–1741 (1999).
- Kwok, S. F. *et al.* Characterization of two subunits of *Arabidopsis* 19S proteasome regulatory complex and its possible interaction with the COP9 complex. *J. Mol. Biol.* **285**, 85–95 (1999).
- Deng, X. W. *et al.* *cop1*: a regulatory locus involved in light-controlled development and gene expression in *Arabidopsis*. *Genes Dev.* **5**, 1172–1182 (1992).
- Kwok, S. F. *et al.* A complement of ten essential and pleiotropic *Arabidopsis* COP/DET/FUS genes is necessary for repression of photomorphogenesis in darkness. *Plant Physiol.* **110**, 731–742 (1996).
- Karniol, B. *et al.* The *Arabidopsis* homologue of an eIF3 complex subunit associates with the COP9 complex. *FEBS Lett.* **439**, 173–179 (1998).
- Staub, J. M. *et al.* Evidence for FUS6 as a component of the nuclear-localized COP9 complex in *Arabidopsis*. *Plant Cell* **8**, 2047–2056 (1996).
- Wei, N. *et al.* *Arabidopsis* COP9 is a component of a novel signaling complex mediating light control of development. *Cell* **78**, 117–124 (1994).
- Serino, G. *et al.* *Arabidopsis* *cop8* and *fus4* mutations define the same locus that encodes subunit 4 of the COP9 signalosome. *Plant Cell* **11**, 1967–1980 (1999).
- von Arnim, A. G. Genetic and developmental control of nuclear accumulation of COP1, a repressor of photomorphogenesis in *Arabidopsis*. *Plant Physiol.* **114**, 779–788 (1997).
- Osterlund, M. T. *et al.* The role of COP1 in repression of *Arabidopsis* photomorphogenic development. *Trends Cell Biol.* **9**, 113–118 (1999).
- Osterlund, M. T. & Deng, X. W. Multiple photoreceptors mediate the light-induced reduction of GUS-COP1 from *Arabidopsis* hypocotyl nuclei. *Plant J.* **16**, 201–208 (1998).
- Reed, J. W. *et al.* Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout *Arabidopsis* development. *Plant Cell* **5**, 147–157 (1993).
- Whitelam, G. C. *et al.* Phytochrome A null mutants of *Arabidopsis* display a wild-type phenotype in white light. *Plant Cell* **5**, 757–768 (1993).
- Ahmad, M. & Cashmore, A. R. HY4 gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366**, 162–166 (1993).
- Lin, C. *et al.* Enhancement of blue-light sensitivity of *Arabidopsis* seedlings by a blue light receptor cryptochrome 2. *Proc. Natl Acad. Sci. USA* **95**, 2686–2690 (1998).

21. Tori, K. U. *et al.* Functional dissection of *Arabidopsis* COP1 reveals specific roles of its three structural modules in light control of seedling development. *EMBO J.* **17**, 5577–5587 (1998).
22. McNellis, T. W. *et al.* Expression of an N terminal fragment of COP1 confers dominant-negative suppression of seedling photomorphogenic development in transgenic *Arabidopsis*. *Plant Cell* **8**, 1491–1503 (1996).
23. Joazeiro, C. A. P. *et al.* The tyrosine kinase negative regulator c-Cbl as a RING-type E2-dependent ubiquitin-protein ligase. *Science* **28**, 309–312 (1999).
24. Loric, K. L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl Acad. Sci. USA* **96**, 11364–11369 (1999).
25. Wei, N. & Deng, X. W. Making sense of the COP9 signalosome. *Trends Genet.* **15**, 98–103 (1999).
26. Reed, J. W. *et al.* Phytochrome A and phytochrome B have overlapping but distinct function in *Arabidopsis* development. *Plant Physiol.* **104**, 1139–1149 (1991).
27. Ahmad, M. *et al.* Cryptochrome blue-light photoreceptors of *Arabidopsis* implicated in phototropism. *Nature* **16**, 720–723 (1994).
28. Wagner, D. *et al.* Overexpression of phytochrome B induces a short hypocotyl phenotype in transgenic *Arabidopsis*. *Plant Cell*, **3**, 1275–1288 (1992).
29. Boylan, M. T. & Quail, P. H. Phytochrome A overexpression inhibits hypocotyl elongation in transgenic *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **88**, 10806–10810 (1991).
30. Koegl, M. *et al.* A novel ubiquitination factor, E4 is involved in multiubiquitin chain assembly. *Cell* **96**, 635–644 (1999).

Acknowledgements

We thank R. Fry and H. Wang for reading and commenting on this manuscript, and A. Cashmore, C. Lin, P. Quail and G. Whitelam for providing *Arabidopsis* photoreceptor mutant strains. Our work was supported by grants from NIH (to X.W.D.), USDA (to N.W.) and the Human Frontiers Science Program Organization. X.W.D. is an NSF Presidential Faculty Fellow, M.T.O. is an NIH and Dept of Education predoctoral trainee and C.S.H. is a Human Frontier Science Program Organization postdoctoral fellow.

Correspondence and requests for materials should be addressed to X.W.D. (e-mail: xingwang.deng@yale.edu).

Genomic rearrangement in *NEMO* impairs NF- κ B activation and is a cause of incontinentia pigmenti

The International Incontinentia Pigmenti (IP) Consortium

* A full list of authors appears at the end of the paper

Familial incontinentia pigmenti (IP; MIM 308310) is a genodermatosis that segregates as an X-linked dominant disorder and is usually lethal prenatally in males. In affected females it causes highly variable abnormalities of the skin, hair, nails, teeth, eyes and central nervous system. The prominent skin signs occur in four classic cutaneous stages: perinatal inflammatory vesicles, verrucous patches, a distinctive pattern of hyperpigmentation and dermal scarring¹. Cells expressing the mutated X chromosome are eliminated selectively around the time of birth, so females with IP exhibit extremely skewed X-inactivation². The reasons for cell death in females and *in utero* lethality in males are unknown. The locus for IP has been linked genetically to the *factor VIII* gene in Xq28 (ref. 3). The gene for *NEMO* (NF- κ B essential modulator)/IKK γ (IKB kinase- γ) has been mapped to a position 200 kilobases proximal to the *factor VIII* locus⁴. *NEMO* is required for the activation of the transcription factor NF- κ B and is therefore central to many immune, inflammatory and apoptotic pathways^{5–9}. Here we show that most cases of IP are due to mutations of this locus and that a new genomic rearrangement accounts for 80% of new mutations. As a consequence, NF- κ B activation is defective in IP cells.

The extreme skewing of X-inactivation observed in IP patients has implications for screening of candidate genes. Complementary DNA prepared from females with IP will not contain the mutated allele of the disease locus, so screening of candidate genes for mutation must be performed with genomic DNA. We used two

approaches to screen *NEMO*: (1) generation and sequencing of fibroblast cDNA from extremely rare cases that express only the mutated X chromosome (two male cases and one female abortus, see Methods); and (2) resolution of the *NEMO* genomic structure for mutational analysis of individual exons. Reverse transcriptase polymerase chain reaction (RT-PCR) from messenger RNA derived from skin fibroblasts from four affected fetuses (D, K, IP85m and G) and controls (C1 and C2) was conducted with *NEMO*-specific primers (Fig. 1). Male samples D and IP85m have only one X chromosome. Female sample K is unusual in that contrary X-inactivation skewing has resulted in expression of only the mutated (IP) X chromosome. Affected female G expresses both Xs, as expected for a fetal IP case that has not yet undergone selective elimination of affected cells.

Amplification between primers located in *NEMO* exons 2 and 3 (exon organization in Fig. 2) for fetuses K, D and G produced a product corresponding to the 5' end of *NEMO* cDNA. However, amplification between exons 2 and 4 produced the predicted fragment from fetus G and control RNAs but not from fetuses D and K. Thus, the 3' end of the *NEMO* cDNA is absent in fetuses K and D. This could not be explained by loss of individual exons because we could amplify all 10 coding exons from genomic DNA (not shown). These results indicate that a mutation that disturbs mRNA production may exist between primers R4 and R1 in exons 3 and 4, respectively.

In a sample from another affected male (IP85m), RT-PCR across the entire coding region of *NEMO* was successful and the products

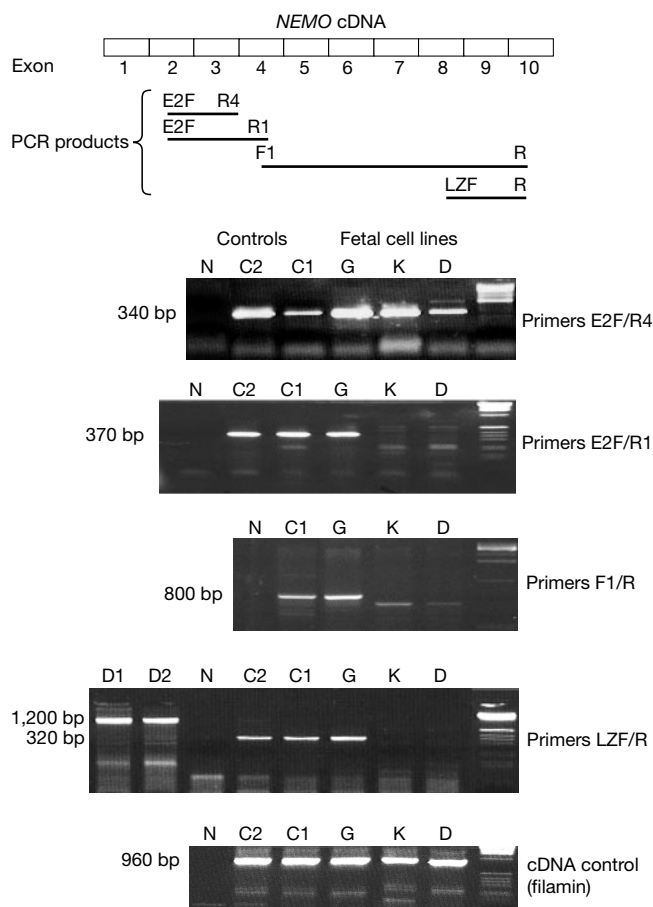


Figure 1 Amplification by RT-PCR of *NEMO* cDNA from IP patient cell lines. The positions of PCR products relative to *NEMO* exons are shown at the top. Amplification is from fetal cDNA samples G, K and D, cDNA from healthy donors (C1 and C2) and control DNA (D1 and D2). PCR of filamin cDNA from all samples acted as a control for RNA integrity. Primer sequences are given in Methods.

sequenced completely. An A-to-G change (base 1,259) was detected within the STOP codon of the mature message (see Supplementary Information). This would result in the addition of 27 residues to the carboxy terminus of the NEMO protein. Examination of the relevant region of genomic DNA (exon 10) from family IP85 revealed that the male proband's affected mother is heterozygous for the change, whereas his unaffected grandmother and grandfather have an intact STOP codon. The mutation has therefore appeared *de novo* with the disease in this family. To determine the spectrum of mutations in additional IP families, we sequenced the complete *NEMO* locus (Fig. 2). The 23-kilobase (kb) gene is composed of 10 exons with three alternative non-coding first exons^{4,5,10} (1a, 1b and 1c; Table 1). The *NEMO* gene partially overlaps the *G6PD* gene and is transcribed in the opposite direction⁴.

Exons 1–9 and the coding portion of exon 10 were amplified by PCR from 30 classically affected, unrelated IP patients (see Table 2 for primer sequences). Amplification products were screened for mutations by single-strand conformation polymorphism (SSCP), heteroduplex analysis or direct sequencing. Only four additional changes were found (Fig. 2 and see Supplementary Information). A 10-base pair (bp) insertion in exon 2 (between coding nucleotides

127 and 128) in family XL349 was found in both an IP proband and her affected mother but not in unaffected siblings. The insertion is an exact duplication of the previous 10 nucleotides, shifts the reading frame of the protein to add eight amino acids after residue 43 and then truncates the protein at the next, in-frame STOP codon. A second insertion in exon 9 of a single C (1110–1111) within a run of cytosines segregates with the disease in family XL203. This frameshift mutation would putatively append 23 new amino acids onto proline residue 370 of the translated protein. A mis-sense transition (A/G) in exon 10 that changes a methionine to a valine in the C terminus of the NEMO protein (M407V) was found in an affected female from family 72. SSCP analysis showed that this relatively conservative change segregates with the disease through a three-generation pedigree. Finally, a C-to-T (C184T), proline to STOP codon mutation found in exon 2 in family XL352 predicts a truncated protein consisting of the amino-terminal 61 amino acids of NEMO. These mutations were not detected in more than 60 unrelated X chromosomes. In addition to apparently pathogenic changes, a polymorphism in exon 1c in the 5' untranslated region (–165 from exon1c and 5,678 of the *NEMO* genomic sequence) changes a C to a G in 6 out of 40 chromosomes.

Surprisingly few mutations were detected by screening individual

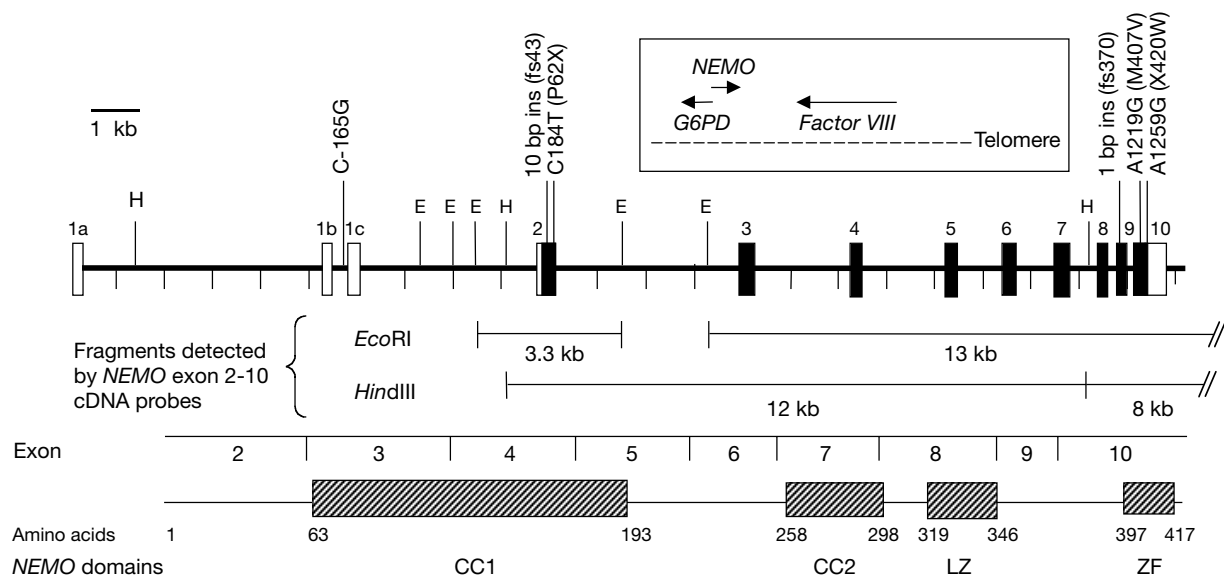


Figure 2 Genomic structure of the *NEMO* gene and partial restriction map (H, *HindIII*; E, *EcoRI*). The 10 exons span 23 kb of genomic sequence. Non-coding (open boxes) and coding exons (filled boxes) are shown. Three alternative first exons numbered 1a, 1b and 1c are found spliced to exon 2. The positions of intragenic mutations, a common polymorphism and restriction fragments seen on Southern blots are shown. The

relationship of coding exons 2–10 to the two coiled-coil (CC1 and CC2), leucine zipper (LZ) and zinc finger (ZF) motifs of the NEMO secondary structure is illustrated. Inset, the orientation of *NEMO* to *factor VIII*, *G6PD* and the Xq28 telomere (not to scale) with arrows in a 5' → 3' direction.

Table 1 Intron-exon structure of the human *NEMO* gene

Intron	Size (bp)	Donor*	Acceptor*
1a	5,263	GTCCATCAGgtggggaag	gaagggcgacCGCGAAACTG
1b	174	CTGCTGACAGgtgtggtcct	acttggggcgCGAGCTGGAC
1c	3,889	TGGCAGCTAGgtatctgatt	tctcttccagCCCTTGCCCT
2	3,981	GAGCTCCGAGgtgaggaaag	ctgccaccagATGCCATCCG
3	2,155	ATGCCAGCAGgtagtgcggg	tatcctgcagCAGATGGCTG
4	1,757	TGGAGGTCGgtgagtcggg	ccgtgccagGGCCCGGGCG
5	1,128	CGGAGGAGAAgtgagtcagc	cttctcagGAGGAAGCTG
6	1,023	CGGAAGCGAgtagtgcca	cgtccttagGGAATGCAGC
7	607	GAAGGCCAGgtgagggccc	gatttgccagCGGATATCT
8	257	AGTCGGCCAGgtgggcctct	gttttcaagGATCGAGGAC
9	298	CCGCCCTGgtgagtgagc	ctttccagCCTACCTCTC

* Upper case denotes exon sequence, lower case denotes intron sequence.

exons of *NEMO* in IP patients. Moreover, amplification and sequencing of all coding exons failed to detect any mutation in fetuses K and D that could explain the abnormal mRNA structure. We therefore investigated the possibility of genomic rearrangement by Southern blotting. Genomic DNAs from IP patients and healthy individuals were digested with *HindIII* and *EcoRI* and hybridized to a probe representing the coding region of *NEMO* cDNA. In addition to the expected wild-type fragments (12-kb and 8-kb *HindIII*, 13-kb and 3.3-kb *EcoRI*), we found a novel 8-kb *HindIII* band and a novel 2.8-kb *EcoRI* band for 38 out of 47 unrelated IP patient samples (Fig. 3a). The novel fragments were concordant for the two enzymes in all cases and were absent for the affected male (IP85m, above) containing a mutation in the termination codon (not shown). The novel bands segregated with disease in each pedigree analysed. In two families an *EcoRI* band of about 14 kb was observed along with the novel 2.8-kb band (for example, IP64, Fig. 3a), with a concomitant reduction in signal for the wild-type 13-kb band. This band did not segregate with disease in these families and is therefore a non-pathogenic variant.

The appearance of novel bands is associated with the development of IP. For two families with sporadic cases (IP44 and IP91) the novel bands (8-kb *HindIII* and 2.8-kb *EcoRI*) were found in the proband's DNAs but not the parents' and have therefore appeared *de novo* with the disease (Fig. 3b). To determine whether the new bands resulted from a deletion, we screened hemizygous DNA from a male abortus whose mother carries the rearrangement (IP1m). This male is from a four-generation IP family (IP1; ref. 11) and has inherited the IP haplotype from his affected mother (data not shown). *NEMO* exons 1a–10 could be amplified by PCR. When DNA from this male was analysed by Southern blotting, 13-kb and 3.3-kb wild-type *EcoRI* bands and the novel 2.8-kb band were seen (lane 12, Fig. 3b). As only one X chromosome is present in this fetus, more than one sequence homologous to the *NEMO* gene must be assumed. The presence of a doublet of hybridizing *HindIII* fragments at about 12 kb in all lanes implies that at least part of *NEMO* is also duplicated in wild-type DNA. Complementary DNA probes encompassing coding sequences from exons 2 and 3 detected the novel IP-specific fragments but probes containing sequences from exons 4–10 did not. Furthermore, hybridization with single exons revealed that exon 3, but not exon 2 or 4, is present on the novel *EcoRI* fragment, whereas exons 2 and 3 but not 4 hybridize to the novel *HindIII* band (not shown).

To investigate whether intronic sequences are involved in the rearrangement, DNA from the male abortus (IP1m) was used to amplify across *NEMO* introns. Amplification across single introns (with forward and reverse primers flanking individual exons; Table 2) was successful for all introns (not shown). However, differences were found compared with controls when PCR reactions were anchored at exon 2 (Fig. 3c). In the affected male (IP1m), amplification could proceed from exon 2 to exon 3 but not from exon 2 to exons 4, 5, 6 or 7. The mutation therefore involves disruption within intron 3. Anchoring the PCR reactions at exons 3 or 7, however, yielded wild-type products for all combinations with downstream exons (Fig. 3c). This is compatible with the existence of a putative second copy of *NEMO* that contains exons 3–10 but not exons 1 and 2. This second copy must be highly homologous: sequencing of long-range PCR products from IP1m revealed that copies of exon 3 in tandem with either exon 2 or exon 4 were identical. Importantly, analysis of DNA from IP-affected male fetus D yielded results similar to those obtained for IP1m for both long-range PCR and Southern blotting, indicating that the common genomic rearrangement is responsible for the production of aberrant *NEMO* mRNA lacking sequences from exons 4–10 as described above. Our RT-PCR, Southern and genomic PCR data indicate that a common rearrangement within intron 3 of *NEMO* generates novel restriction fragments that contain exon 3.

We used this information to isolate the boundary of rearrangement neighbouring exon 3 within the 2.8-kb novel *EcoRI* band. DNA from the male abortus (IP1m) was digested with *EcoRI* and ligated to pre-annealed adapter oligonucleotides with *EcoRI* cohesive ends (see Methods). Nested PCR between adapter primers and exon-3-specific forward primers yielded a product of about 2 kb that was sequenced completely. The fragment was found to extend from exon 3 into intron 3 for 1.8 kb. The sequence then diverged from intron 3 identity after a position equivalent to 15,749 of the *NEMO* genomic sequence; a novel 104-bp sequence was found between this boundary and the *EcoRI* site. Interestingly, the boundary between intron 3 and this sequence coincides precisely with the end of a repeat of the MER67B family in intron 3. The 104-bp sequence, which itself is homologous to an L2 repeat, was used to screen available sequence data from Xq28 BAC clones and found within a short sequence contig from BAC clone 211L10, which also contains the *NEMO* gene (for sequence see Supplementary Information).

Table 2 Primers and conditions for PCR amplification of individual exons

Exon	Length (bp) Product size (bp)	Forward primer (5' → 3')	Reverse primer (5' → 3')
1a	125 302	GGA AGT CAG CCC AGA AAT GT	GTA CTT ACT GCC CCC TTC CA
1b	95 297	GGG CGG GGC TTG TGT TTT TA	AGA AGC GCG GAG CAG GAA CG
1c	170 334	TTC CTG CTC CGC GCT TCT GG	CCA GGA GAG CCC ATT CAT TC
2	202 321	TCT GCT GGG TAA GGA TGT GG	TCT GCA GGT GGG GAG AAG AC
3	212 296	CCC AGC TCC CCT CCA CTG TC	CAC CTG GCG TCA CTC GGC GGG T
4	119 194	CAG TGC TGA CAG GAA GTG GC	AAC CCT GGA AGG GGT CTC CGG AG
5	153 351	CAT CAG CTC GCA GTC ACA GG	CCG ACA CTT CTC AGC CTT TC
6	97 277	AAG GGG GTA GAG TTG GAA GC	AGG CAA GTC TAA GGC AGG TC
7	144 347	GCC ACT CTT TCA TCC TTC TC	CTG GGC AAC AAG AGC AAA AC
8	143 245	TGC CTG GTG GGT GGC TGG CT	CAG TGT CGC ACC CAC TGC TCA
9	62 234	GCT GCT TTG ACA CTA GTC CA	CAG AGA GCA ACA GGA AGG TC
10	728 363	CGG CGG CTC CTG GTC TTA CA	GCC ACC CAG CCC TTC ATC CT

To confirm that the novel sequence is associated with the rearranged 2.8-kb *EcoRI* fragment, we conducted PCR between exon 3 and the novel boundary sequence for families with and without the common rearrangement found on Southern blots (Fig. 3d). The resulting 2-kb band occurs only in IP patients who have been shown to carry the altered band by Southern blotting, not in controls and parents of sporadic cases, and is a simple diagnostic PCR test for 80% of IP cases. The presence of a 13-kb band in patients or controls without the rearrangement suggested that the natural site for the boundary sequence may be 3' of *NEMO*. End sequencing of both this band and a 13-kb *EcoRI* fragment housing exons 3–10 of *NEMO* from BAC clone 211L10 confirmed this hypothesis. Furthermore, at this downstream location sequences identical to the MER67B of intron 3 were found juxtaposed to the 104 bp of L2 repeat in the same orientation as in the *NEMO* gene. Further PCR and sequencing indicated that although 870 bp of intron 3 is repeated 3' of the *NEMO* gene, there is no additional exon 3 within this fragment.

The most common mutation in IP is therefore a deletion of part of the *NEMO* gene mediated by directly repeated sequences within intron 3 and 3' of exon 10 (Fig. 3e). The translocation of sequences

from this downstream region to intron 3 can explain the novel band sizes observed on Southern blots, as *EcoRI* and *HindIII* sites are found at appropriate 3' positions. That the rearrangement gives rise to aberrant *NEMO* message has also been proved by RT-PCR amplification between L2 primer JF3R and exon 3 forward primers on cDNA from fetus D and K (not shown). This revealed that at least 26 novel amino acids would be appended to the 133 residues of *NEMO* encoded by exons 2 and 3. A second, homologous but incomplete copy of the *NEMO* gene, suggested by Southern and PCR data, appears to be unaltered by the IP mutation (Fig. 3e). IP therefore joins a growing list of disorders caused most frequently by genomic rearrangement¹². The origin of the genomic rearrangements often shows a predilection for one parent. For example, homologous recombination between tandem repeats leading to *de novo* duplication in CMT1A occurs primarily during male gametogenesis¹³. For IP, two-thirds of new mutations originate from fathers, consistent with paternal bias for rearrangement². The origin of mutation has been established for 12 patients with the rearrangement. Ten occurred during male gametogenesis, implicating intrachromosomal interchange. In cases of genomic rearrangement, it is important to ascertain whether more than one

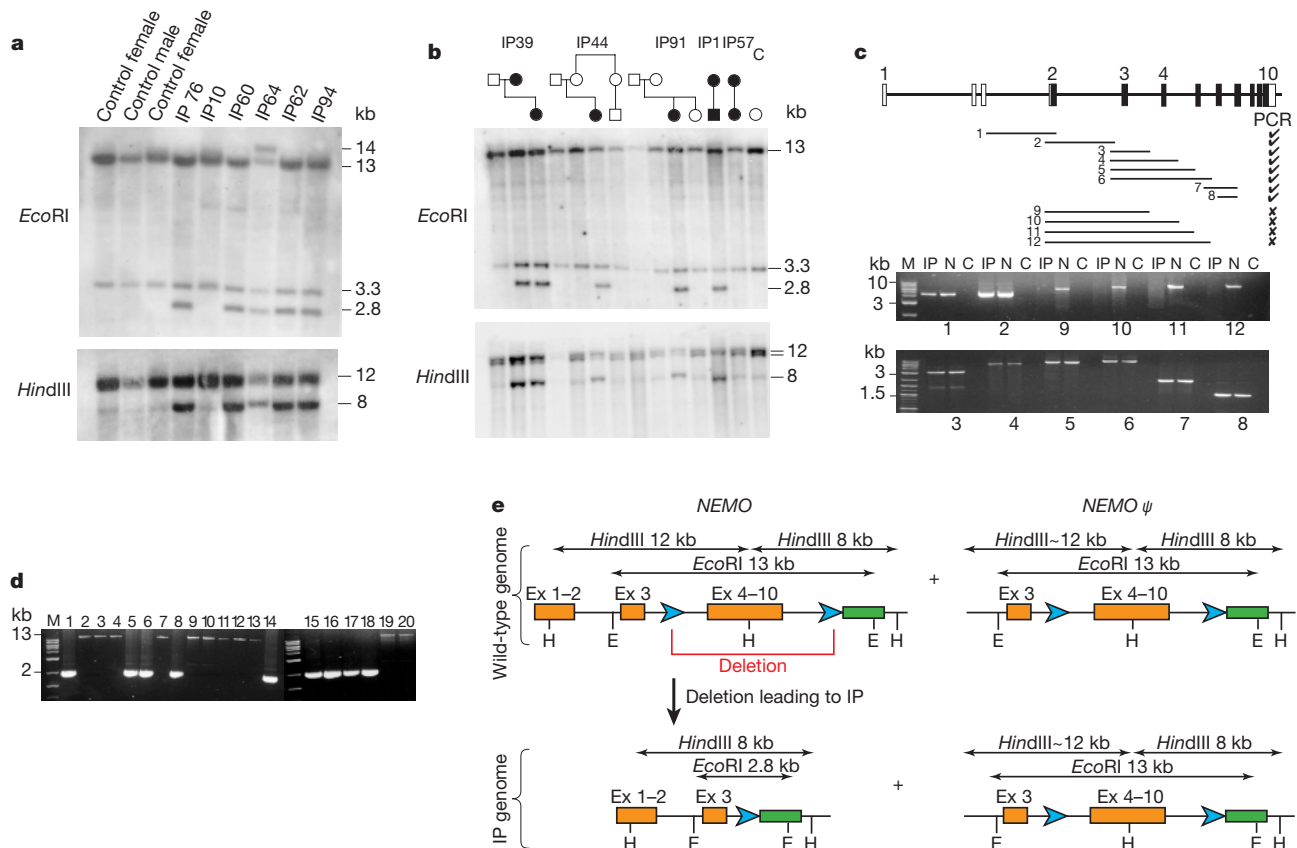


Figure 3 Genomic rearrangement in IP patients. **a**, Southern blot hybridized with *NEMO* cDNA representing exons 2–10 showing the appearance of new *EcoRI* (2.8 kb) and *HindIII* (8 kb) bands in unrelated IP females compared with controls. IP10 does not show these bands for either enzyme. A female in family IP64 also has a novel 14-kb band that does not segregate with disease. For *HindIII*, a doublet of bands at 12 kb is seen instead of the expected single band in addition to bands corresponding to exons 7–10 (faint 8-kb band). **b**, Novel *EcoRI* and *HindIII* bands appear *de novo* with the disease in sporadic cases of IP (IP44 and IP91). Segregation in family 39 is also shown. A male sample from family IP1 carrying the rearrangement still retains wild-type *EcoRI* bands although the smaller of the doublet *HindIII* bands has disappeared. C, control female DNA. **c**, Long-range PCR across the *NEMO* gene in male patient IP1m. Schematic shows 12 PCR reactions conducted in patient (IP) versus control (N) samples, and the gel shows the products. Ticks, successful

amplification in IP1m; crosses, lack of amplification. **d**, Diagnostic PCR in IP patients with the rearrangement. Amplification between exon 3 forward primer 3F8 and a primer from the L2 repeat in the rearranged *EcoRI* fragment (JF3R) gives a 2-kb band only in IP samples (lanes 1, 5, 6, 8, 14–18). A 13-kb band is seen in control samples that do not have the rearrangement (lanes 2–4, 7, 9–13, 19, 20). **e**, Model for IP rearrangements. In addition to the *NEMO* gene (left), a *NEMO* pseudogene lacking exons 1 and 2 is proposed (right, *NEMO*ψ). An 870-bp region of identity corresponding to a MER67B repeat (arrow) exists both in intron 3 and 3' to exon 10. An L2 repeat follows the 3' MER67B sequence at the downstream site (hatched box). Recombination between the regions of identity will delete exons 4–10 of *NEMO*. Diagnostic PCR between exon 3 and the L2 sequence yields a 2-kb band in patients and a 13-kb band in controls.

gene is involved. Of the nine out of 47 IP patients found not to have a rearrangement, six have been screened for intragenic changes and four found to have mutations (above). These segregate with the disease in families or have arisen *de novo* with the disease, indicating that defects in *NEMO* alone are sufficient to cause the disorder.

NEMO is essential in the NF- κ B activation process⁹. NF- κ B homo- or heterodimers are sequestered in the cytoplasm through interaction with an inhibitory molecule of the I κ B family. Upon cytokine stimulation, the I κ B molecules are phosphorylated, poly-ubiquitinated and degraded through the ubiquitin-proteasome pathway^{6,14}. NF- κ B is then free to translocate to the nucleus and to activate its target genes. This phosphorylation event is carried out by a high molecular mass, multiprotein kinase complex containing two subunits with kinase activity (IKK1/ α and IKK2/ β). The third known component of this IKK complex is *NEMO* (IKK γ , IKKAP or IKBKG Human Gene Nomenclature)^{5,7,9}, a protein with relative molecular mass 48,000 (M_r 48K). *NEMO* has no apparent catalytic activity; it interacts directly with the kinase subunits and is required for activation of the kinase complex in response to extracellular (or intracellular) stimuli. Its absence results in a complete inhibition of NF- κ B activation. We therefore determined whether NF- κ B activation was defective in IP patients K and D, who express only mutated

NEMO, compared with controls. These patients possess the common rearrangement found in 80% of IP cases.

We analysed cytoplasmic extracts from primary embryonic fibroblasts by western blotting for the abundance of the three known subunits of IKK (IKK-1, IKK-2 and *NEMO*), the two major subunits of NF- κ B (p50 and relA) and two I κ B species (I κ B α and β) (Fig. 4a). No 48K band corresponding to the *NEMO* molecule was detected in fibroblasts from patients K and D compared to controls, although a low relative molecular mass band of about 21K was seen. This may represent the truncated version of *NEMO* predicted from genomic and RT-PCR analysis. Analysis of IKK-1, IKK-2, p50 and relA protein levels did not reveal any significant differences between patients and controls. However, a substantial increase in I κ B α and I κ B β was seen in fibroblasts from patients K and D. As the stability and levels of I κ Bs are controlled by NF- κ B, this observation indicated that the NF- κ B system may be perturbed in IP patients.

NF- κ B activation was measured in embryonic fibroblasts from patients K and D with an electrophoretic mobility shift assay (Fig. 4b). Upon stimulation with tumour necrosis factor (TNF), two retarded complexes were observed with nuclear extracts prepared from normal and patient G embryonic fibroblasts. A supershift experiment showed that the upper complex was composed of both

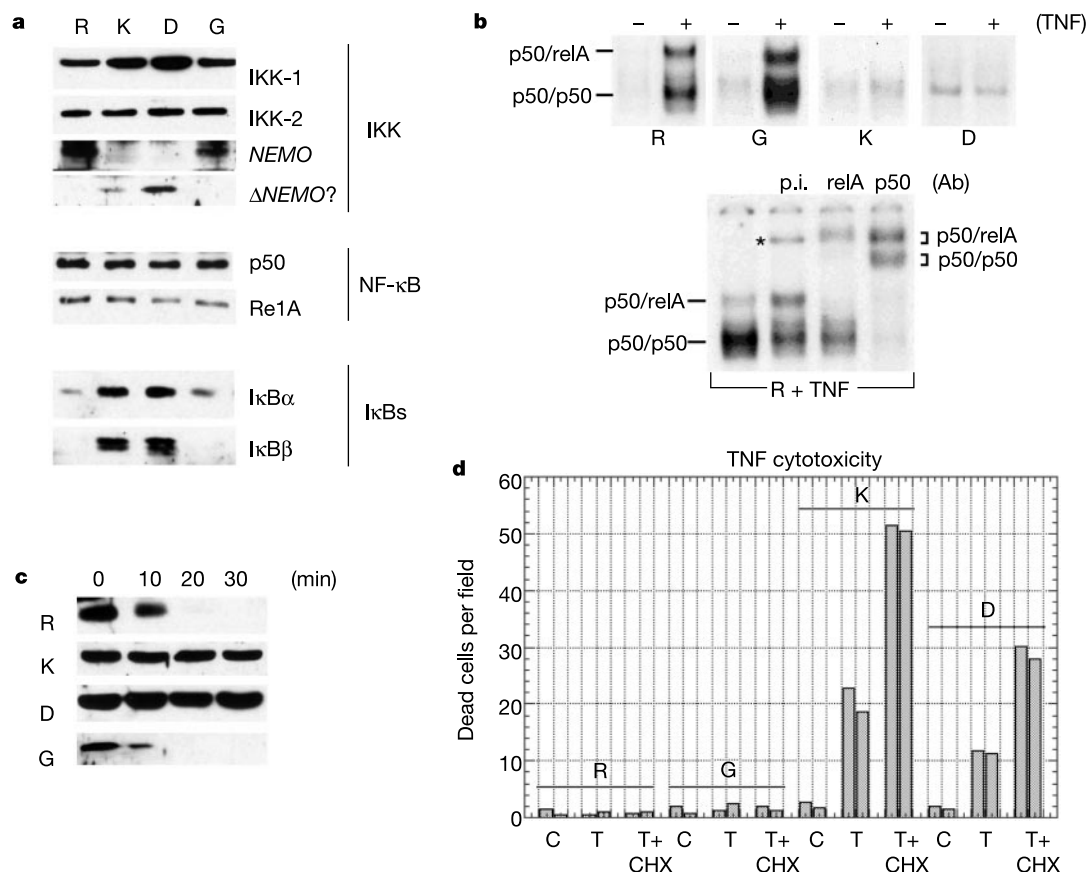


Figure 4 Defective NF- κ B activation in IP cells. **a**, *NEMO* is not detected in IP embryonic fibroblasts. Cytoplasmic extracts derived from normal, G, K or D embryonic fibroblasts were subject to western blotting with antibodies directed against *NEMO*, IKK-1, IKK-2, p50, relA, I κ B α and I κ B β . Δ *NEMO*, putative truncated version of *NEMO* found only in patients K and D. **b**, NF- κ B is not activated in IP embryonic fibroblasts. Top, normal fibroblasts or those derived from patients G, K or D were mock- or TNF α -treated. Retarded complexes corresponding to p50/relA and p50/p50 dimers (see below) are indicated. Bottom, nuclear extracts prepared from TNF-treated normal embryonic fibroblasts were pre-incubated with pre-immune serum (p.i.) or with anti-relA or p50 sera, then analysed

with an EMSA. Asterisk, nonspecific band generated with the p.i. serum. Supershifted p50/relA and p50/p50 dimers are indicated. **c**, I κ B α degradation is defective in IP embryonic fibroblasts. Normal cells or those derived from patients G, K or D were treated with TNF α (10 ng ml⁻¹) for the indicated times, and cytoplasmic extracts were prepared and analysed, after western blotting, with anti-I κ B α antibody. **d**, IP embryonic fibroblasts are TNF-sensitive. Normal cells or those derived from patients G, K or D were mock (C), TNF α (T) or TNF α plus cycloheximide (T+CHX)-treated for 20 h, and the cell viability was determined through trypan blue exclusion. One experiment, out of two giving similar results, is shown.

p50 and relA proteins, whereas the lower one was composed of p50 only (Fig. 4b). Importantly, neither p50/relA nor p50/p50 complex was induced in nuclear extracts derived from TNF-treated embryonic fibroblasts from patients K and D, indicating a complete defect in the NF- κ B activation process. A similar observation was made when interleukin-1 (IL-1) was used instead of TNF (data not shown).

As the translocation of NF- κ B to the nucleus results from degradation of I κ B, the fate of I κ B α in normal and IP embryonic fibroblasts after treatment with TNF or IL-1 was determined. Whereas patient G and control fibroblasts exhibited complete disappearance of I κ B α after 40 min of TNF or IL-1 treatment, no such disappearance could be observed in cells from patients K and D (Fig. 4c, and data not shown).

In several cell types, including embryonic fibroblasts, NF- κ B protects against TNF-induced apoptosis¹⁵. We investigated the sensitivity of normal and IP embryonic fibroblasts to TNF (Fig. 4d). In contrast to cells from controls or patient G, which were unaffected by treatment with TNF or TNF and cycloheximide, embryonic fibroblasts from patients K and D were sensitive to TNF and this sensitivity was increased by cycloheximide. These data demonstrate a lack of NF- κ B activation in IP embryonic fibroblasts, resulting from a defect at the level of the NEMO molecule. As a consequence, IP cells are highly sensitive to pro-apoptotic signals.

Can the pathology observed in male and female IP patients be explained in terms of NEMO function? The genes encoding components of the IKK complex have been inactivated by homologous recombination in the mouse. Inactivation of IKK2 resulted in embryonic death due to massive liver apoptosis at day 14 (for a review see ref. 6), whereas that of NEMO resulted in death of male embryos at day 12 with a similar phenotype¹⁶ (murine NEMO is also X-linked). A similarly marked phenotype was observed when the gene encoding relA, the most potent transcriptional activator of the NF- κ B family, was inactivated¹⁷. This phenotype was due to the pro-apoptotic effect of TNF¹⁸, in keeping with the high sensitivity to TNF-induced apoptosis observed in cell lines derived from IP patients. In both mice and humans, therefore, the complete absence, or strong reduction, of NF- κ B activity results in embryonic lethality.

A prominent role for NF- κ B in the apoptotic response also has a bearing on the phenotype in heterozygous females. It is likely that skewing of X-inactivation, at least in blood, results from progressive apoptotic elimination of cells bearing mutated *NEMO* on the active X chromosome. For NEMO-deficient heterozygous female mice, a phenotype comparable to that seen in man has not yet been reported. Owing to lyonization, however, understanding the effects observed in human patients is difficult. Further clues emerge from targeting of other components of the NF- κ B pathway in mouse¹⁹. The individual inactivation of four of the five known members of the NF- κ B family results in more or less severe defects in the immune response. More interestingly, mice devoid of both the p105/p50 and the p100/p52 subunits fail to generate mature osteoclasts, causing severe osteopetrosis^{20,21}. Osteoclasts are also essential for tooth eruption as they resorb the alveolar bone to form an eruption pathway, and eruption abnormalities are a prominent feature of the human IP phenotype. Mice carrying inactivation of IKK1 (refs 22–24) exhibit almost intact activation of NF- κ B by pro-inflammatory stimuli, but show defects in morphogenetic events, including limb and skeletal patterning, and proliferation and differentiation of epidermal keratinocytes. At the skin level, NF- κ B appears to have a dual role: it controls cell growth in the stratified epithelium and regulates apoptosis. Defects in both pathways may explain the characteristic skin lesions observed in IP. Also reminiscent of the abnormalities observed in female IP patients are the deformed incisors and lack of hair follicles in the IKK1^{-/-} mice.

The phenotype observed in a surviving male with a mutation in NEMO (IP85m) provides a unique view of the consequences of impaired NF- κ B function in man. Severely compromised immune

cell function, lack of teeth and osteopetrosis (see Methods) are consistent with mouse models of defective NF- κ B function. A comparison can be made between the phenotype in IP85m and that of patients with familial expansile osteolysis (FEO, or familial Paget disease of bone, PDB). This is an autosomal dominant disorder characterized by focal areas of increased bone remodelling. In contrast to our findings, activating mutations in the gene for a positive activator of NF- κ B function (RANK) have been found to account for this disorder in two families²⁵. Interestingly, IP85m also showed signs of capillary bed abnormalities, both in the skin and the gut, suggesting an unsuspected role for NF- κ B in the maintenance of blood vessel architecture.

The chromosomal rearrangement found in most patients would result in a truncated molecule carrying the 133 N-terminal amino acids of NEMO plus an unknown number of novel residues. This molecule would contain part of the first coiled-coil domain (Fig. 2), and may still interact with IKK-2 but is unlikely to respond to upstream signals. The mutation in family XL349 results in a NEMO molecule consisting of only 61 N-terminal amino acids. This short peptide could not interact with IKKs and is probably inactive. The insertion that introduces a frameshift after amino acid 370 results in the replacement of the entire C-terminal, zinc finger domain with 23 unrelated amino acids. Deletion of this zinc finger domain results in a severe loss of function of the NEMO molecule in complementation assays (G.C. and A.I., unpublished data). Thus, with rare exceptions, mutations causing IP severely truncate the NEMO protein and result in loss of function.

An interesting mutation is that affecting the STOP codon in male patient IP85m. A likely explanation for this milder phenotype is partial loss of NEMO function, raising the possibility that mis-sense mutations with only minor effects on NEMO function may cause a less morbid phenotype. Therefore, mis-sense mutations in *NEMO* may contribute to a wide range of conditions related to IP such as immune deficiency, osteopetrosis, and dental and ophthalmological abnormalities. □

Methods

Collection of families and preparation of samples

Families with IP were ascertained with ethical committee approval at the collaborating institutions. All affected females had a history of perinatal blistering and at least one other stage of skin lesions, dental, and/or eye, and/or central nervous system abnormalities.

Fibroblast lines

Primary fibroblast cultures expressing a mutant IP X chromosome were available from four subjects. Fetus K was a spontaneously aborted, affected female belonging to a three-generation IP family. Lethality in this case was associated with exclusive expression of the maternal X chromosome bearing the IP mutation, presumably because a deleterious mutation is present on the opposing X or for stochastic reasons (A.S. *et al.*; manuscript in preparation). Fetus D was a spontaneously aborted, affected male with an affected mother. Fetus G was a medically interrupted affected female belonging to a large IP family. X-inactivation was random in this line. IP85m was the affected son of a classically affected IP female. He was born with multiple capillary hemangiomas, developed lymphedema of the lower limbs and failed to thrive owing to malabsorption. Despite a destructive red blood cell picture and recurrent infections due to poor immune cell function, he survived two and a half years then succumbed to a tuberculosis infection. He had operations to remove his spleen and a gut stricture, and biopsies revealed abnormal capillary beds in the gut, extrahepatic erythropoiesis and osteopetrosis. His skin developed a reticular pigmentation. His cognitive development was normal.

DNA technology and mutation screening

DNA was extracted from fibroblasts or lymphoblasts. Southern blots were conducted with neutral transfer onto Immobilon NY+ (Millipore) and ultraviolet crosslinking. PCR amplification of individual NEMO exons was conducted with Amplitaq gold and primers described in Table 2. Long-range PCR was conducted using EXPAND long template PCR system (Roche Molecular Biochemicals). DNA sequencing was conducted with BigDye terminator cycle sequencing (Perkin Elmer). SSCP and Heteroduplex analyses were done as described^{26,27}.

RNA isolation and RT-PCR

Poly A+ RNA was isolated from fibroblast cultures with Ingenious Mini Message Maker (R&D Systems) and cDNA prepared with the AmpGold RNA PCR system (Perkin Elmer) and random primers. Primers used for amplification of NEMO cDNA:

E2F(CCTTGGCCTGTTGGATGAAT), R(CGTCCTCGGCAGGCAGTGGCC), F1(GGCCACTGCGTCCGAGGACG), R1(ACCTCCAGAGCTGGCATTG), R4(CTTCAGATCGAGCTTCTCGAG), LZ(F(GCGGACTTCCAGGCTGAGAGG).

Genomic sequencing of the *NEMO* gene

The *NEMO* gene was contained within BAC clone 211L10 from RPC111 (human male BAC library) (S.A. *et al.*, manuscript in preparation). The gene sequence was determined by a combined strategy of shotgun sequencing of M13 subclones and long-range genomic PCR products. BAC DNA was sonicated and the ends repaired with T4 DNA polymerase. Fragments of 1–2 kb were fractionated from the mixture by agarose gel electrophoresis and subcloned into M13mp18 vector prepared by digestion with *Sma*I. We screened 20,000 clones and sequenced 120 *NEMO*-positive M13 clones. Sequence tracts were assembled with Applied Biosystem's FACTURA and INHERIT programs. Gaps were then closed with PCR to amplify the intervening material. The exon–intron boundaries are given in Table 1. The gene is listed under accession number AJ271718.

Adapter PCR of the intron 3 rearrangement boundary

DNA from a male abortus with the *NEMO* rearrangement from family IP1 (ref. 11) was digested with *Eco*RI and purified by ethanol precipitation. We pre-annealed 750 pmol of adapter primers *Eco* adapt1 (CTAATACGACTCACTATAGGGCTCGAGCAGCTCC-GAGGGCAG) and *Eco* adapt2 (P-AATTCTGCGCCTCGGAG) by denaturing in 30 µl buffer (5 mM Tris pH 7.4, 50 mM MgCl₂). Genomic DNA (100 ng) was ligated to a 250-fold molar excess of pre-annealed adapter primers in a 10 µl ligation reaction with T4 DNA ligase and manufacturers buffer (NEB). Primers AP1 (CCATCCTAATACGACTCACTA-TAGGGC) and intron2/exon 3 forward primer 3F (CCCAGCTCCCCTCCACTGTGTC) were incorporated into a primary PCR reaction with the EXPAND long template PCR system (Roche Molecular Biochemicals). A nested reaction with adapter primer AP2 (TCAC-TATAGGGCTCGAGCAGC) combined with a primer F3 from within exon 3 (CGGCA-GAGCAACCAGATTCTGC) yielded a 2-kb fragment. Diagnostic PCR across the boundary in IP patients was performed with 3FH (GACCAGTCCCCTCCACTGTGTC) forward primer, JF3R (CTCGGAGACAGGAACGAGCA) reverse primer and system 3 of the EXPAND kit (annealing temperature 65 °C, 10 cycles of 7-min extension and 20 cycles with increments of 20 s per cycle).

NF-κB function in IP fibroblast lines

Recombinant human interleukin-1α was from Biogen. Recombinant human TNFα was from PreproTech. Cycloheximide was from Sigma. Antisera: anti-NEMO⁹ anti-p50 1141 and anti-relA 1226 (ref. 28), anti-IκBβ 30715 (ref. 29). Purified polyclonal antibody directed against IKK-2 and IκBα were from Santa Cruz. IKK-1 monoclonal antibody was from Pharmingen. Preparation of cell extracts, western blot analysis and electrophoretic mobility shift assays were performed as described³⁰. For cytotoxicity assays, cells were treated for 20 h in complete growth medium with 20 ng ml⁻¹ TNFα or 20 ng ml⁻¹ TNFα plus 300 ng ml⁻¹ cycloheximide. Viability of cells was estimated by trypan blue exclusion.

1. Landy, S. J. & Donnai, D. Incontinentia pigmenti (Bloch–Sulzberger syndrome). *J. Med. Genet.* **30**, 53–59 (1993).
2. Parrish, J. E., Scheuerle, A. E., Lewis, R. A., Levy, M. L. & Nelson, D. L. Selection against mutant alleles in blood leukocytes is a consistent feature in Incontinentia Pigmenti type 2. *Hum. Mol. Genet.* **5**, 1777–1783 (1996).
3. Smahi, A. *et al.* The gene for the familial form of incontinentia pigmenti (IP2) maps to the distal part of Xq28. *Hum. Mol. Genet.* **3**, 273–278 (1994).
4. Jin, D. Y. & Jeang, K. T. Isolation of full-length cDNA and chromosomal localization of human NF-kappaB modulator NEMO to Xq28. *J. Biomed. Sci.* **6**, 115–120 (1999).
5. Rothwarf, D. M., Zandi, E., Natoli, G. & Karin, M. IKK-γ is an essential regulatory subunit of the IκB kinase complex. *Nature* **395**, 297–300 (1998).
6. Rothwarf, D. & Karin, M. The NF-κB activation pathway: a paradigm in information transfer from membrane to nucleus. *Science's Signal Transduction Knowledge Environment* (cited 26 October 1999) (http://www.stke.org/cgi/content/full/OC_sigtrans;1999/5/re1) (1999).
7. Mercurio, F. *et al.* IκB kinase (IKK)-associated protein 1, a common component of the heterogeneous IKK complex. *Mol. Cell. Biol.* **19**, 1526–1538 (1999).
8. Israël, A. The IKK complex: an integrator of all signals that activate NF-κB? *Trends Cell Biol.* **10**, 129–133 (2000).
9. Yamaoka, S. *et al.* Complementation cloning of NEMO, a component of the IκB kinase complex essential for NF-kappaB activation. *Cell* **93**, 1231–1240 (1998).
10. Li, Y. *et al.* Identification of a cell protein (FIP-3) as a modulator of NF-kappaB activity and as a target of an adenovirus inhibitor of tumor necrosis factor alpha-induced apoptosis. *Proc. Natl Acad. Sci. USA* **96**, 1042–1047 (1999).
11. Jouet, M. *et al.* Linkage analysis in 16 families with incontinentia pigmenti. *Eur. J. Hum. Genet.* **5**, 168–170 (1997).
12. Lupski, J. R. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.* **14**, 417–422 (1998).
13. Lopes, J. *et al.* Sex-dependent rearrangements resulting in CMT1A and HNPP. *Nature Genet.* **17**, 136–137 (1997).
14. Ghosh, S., May, M. J. & Kopp, E. B. NF-κB and rel proteins: Evolutionary conserved mediators of immune responses. *Annu. Rev. Immunol.* **16**, 225–260 (1998).
15. Van Antwerp, D. J., Martin, S. J., Kafri, T., Green, D. R. & Verma, I. M. Suppression of TNF-α-induced apoptosis by NF-κB. *Science* **274**, 787–789 (1996).
16. Rudolph, D. *et al.* Severe liver degeneration and lack of NF-κB activation in NEMO/IKKγ-deficient mice. *Genes Dev.* **14**, 854–862 (2000).
17. Beg, A. A., Sha, W. C., Bronson, R. T., Ghosh, S. & Baltimore, D. Embryonic lethality and liver degeneration in mice lacking the RelA component of NF-κB. *Nature* **376**, 167–170 (1995).

18. Doi, T. S. *et al.* Absence of tumor necrosis factor rescues RelA-deficient mice from embryonic lethality. *Proc. Natl Acad. Sci. USA* **96**, 2994–2999 (1999).
19. Attar, R. M. *et al.* Genetic approaches to study rel/NF-κB/1κB function in mice. *Semin. Cancer Biol.* **8**, 93–101 (1997).
20. Franzoso, G. *et al.* Requirement for NF-κB in osteoclast and B-cell development. *Genes Dev.* **11**, 3482–3496 (1997).
21. Iotsova, V. *et al.* Osteopetrosis in mice lacking nf-kappa-b1 and nf-kappa-b2. *Nature Med.* **3**, 1285–1289 (1997).
22. Hu, Y. L. *et al.* Abnormal morphogenesis but intact IKK activation in mice lacking the IKKα subunit of IκB kinase. *Science* **284**, 316–320 (1999).
23. Li, Q. T. *et al.* IKK1-deficient mice exhibit abnormal development of skin and skeleton. *Genes Dev.* **13**, 1322–1328 (1999).
24. Takeda, K. *et al.* Limb and skin abnormalities in mice lacking IKKα. *Science* **284**, 313–316 (1999).
25. Hughes, A. E. *et al.* Mutations in TNFRSF11A, affecting the signal peptide of RANK, cause familial expansile osteolysis. *Nature Genet.* **24**, 45–48 (2000).
26. Alonso, A., Martin, P., Albarran, C., Garcia, O. & Sancho, M. Rapid detection of sequence polymorphisms in the human mitochondrial DNA control region by polymerase chain reaction and single-strand conformation analysis in mutation detection enhancement gels. *Electrophoresis* **17**, 1299–1301 (1996).
27. Ganguly, A., Rock, M. J. & Prockop, D. Conformation sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. *Proc. Natl Acad. Sci. USA* **90**, 10325–10329 (1993).
28. Rice, N. R., MacKichan, M. L. & Israël, A. The precursor of NF-κB p50 has IκB-like functions. *Cell* **71**, 243–253 (1992).
29. Weil, R., Laurent-Winter, C. & Israël, A. Regulation of IκBβ degradation—similarities to and differences from IκBα. *J. Biol. Chem.* **272**, 9942–9949 (1997).
30. Courtois, G., Whiteside, S. T., Sibley, C. H. & Israël, A. Characterization of a mutant cell line that does not activate NF-κB in response to multiple stimuli. *Mol. Cell. Biol.* **17**, 1441–1449 (1997).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as a paper copy from the London editorial office of Nature.

Acknowledgements

The IP Consortium thanks S. Emmerich, Director of the National Incontinentia Pigmenti Foundation, for financial and practical support, and for fostering a collaborative spirit throughout. We also thank the participating families for their cooperation, and their attending physicians for sharing information and facilitating enrolments: in particular V. Murday, S. Landy, J. Dean, M. C. Hors-Cayla, N. Dahl and M. Gonzales. This work has been supported by Action Research UK (S.J.K.); Association Française contre les Myodystrophies (A.M.); The Foundation Fighting Blindness; Research to Prevent Blindness, Inc.; (R.A.L.); the NIH (DLN); Telethon, Italy (M.D.U); Deutsche Forschungsgemeinschaft; the European Community (EC) Fifth Framework Programme (A.P.); Association pour la Recherche sur le Cancer, Ligue Nationale contre le Cancer and EC (A.I.) and by the National Incontinentia Pigmenti Foundation. R.A.L. is a Research to Prevent Blindness Senior Scientific Investigator. A.I. and G.C. thank C. Bessia for technical help.

Correspondence and requests for materials should be addressed to S.J.K. (e-mail: SJK12@mole.bio.cam.ac.uk).

* The International Incontinentia Pigmenti (IP) Consortium:

France: Asmae Smahi[†], G. Courtois[‡], P. Vabres^{*}, S. Yamaoka[‡], S. Heuertz^{*}, A. Munnich^{*} & A. Israël[‡]

Germany: Nina S. Heiss[†], S. M. Klaubek[§], P. Kioschis[§], S. Wiemann[§] & A. Poustka[§]

Italy: Teresa Esposito[†], T. Bardaro^{||}, F. Gianfrancesco^{||}, A. Ciccodicola^{||} & M. D'Urso^{||}

UK: Hayley Woffendin[†], T. Jakins[†], D. Donnai[#], H. Stewart[#] & S. J. Kenwright[†]

USA: Swaroop Aradhya[†], T. Yamagata^{**}, M. Levy^{††}, R. A. Lewis^{***} & D. L. Nelson[†]

* Department of Genetics, Unité de Recherches sur les Handicaps Génétiques de l'Enfant INSERMU-393, Hôpital Necker-Enfants Malades, 75015 Paris, France
‡ Unité de Biologie Moléculaire de l'Expression Génique URA 1773 CNRS, Institut Pasteur, 75724 Paris cedex 15, France

§ Deutsches Krebsforschungszentrum (DKFZ), Department of Molecular Genome Analysis, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

|| International Institute of Genetics and Biophysics, C.N.R. Via G. Marconi, 10, 80125 Naples, Italy

† Wellcome Trust Centre for Molecular Mechanisms of Disease and University of Cambridge Department of Medicine, Addenbrooke's Hospital, Hill Road, Cambridge CB2 2XY, UK

St Mary's Hospital, Whitworth Park, Manchester M13 0JH, UK
Departments of *Molecular and Human Genetics, †† Dermatology and

** Ophthalmology, Baylor College of Medicine, Houston, Texas 77030, USA

† These authors contributed equally to this work

Functional link between ataxia-telangiectasia and Nijmegen breakage syndrome gene products

Song Zhao^{††}, Yi-Chinn Weng^{††}, Shyng-Shiou F. Yuan^{††‡}, Yi-Tzu Lin^{††}, Hao-Chi Hsu^{*}, Suh-Chin J. Lin^{*}, Elvira Gerbino^{*}, Mei-hua Song^{*}, Małgorzata Z. Zdzienicka[§], Richard A. Gatti^{||}, Jerry W. Shay[¶], Yael Ziv[#], Yosef Shiloh[#] & Eva Y.-H. P. Lee^{*}

^{*} Department of Molecular Medicine/Institute of Biotechnology, The University of Texas Health Science Center at San Antonio, San Antonio, Texas 78245-3207, USA

[§] MGC-Department of Radiation Genetics and Chemical Mutagenesis, Leiden University, LUMC, Leiden, The Netherlands

^{||} Department of Pathology, University of California Los Angeles, Los Angeles, California 90095, USA

[¶] Department of Cell Biology, The University of Texas Southwestern Medical Center, Dallas, Texas 75390-9039, USA

[#] Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

^{††} These authors contributed equally to this work

Ataxia-telangiectasia (A-T) and Nijmegen breakage syndrome (NBS) are recessive genetic disorders with susceptibility to cancer and similar cellular phenotypes¹. The protein product of the gene responsible for A-T, designated ATM, is a member of a family of kinases characterized by a carboxy-terminal phosphatidylinositol 3-kinase-like domain^{2,3}. The NBS1 protein is specifically mutated in patients with Nijmegen breakage syndrome and forms a complex with the DNA repair proteins Rad50 and Mre11⁴⁻⁷. Here we show that phosphorylation of NBS1, induced by ionizing radiation, requires catalytically active ATM. Complexes containing ATM and NBS1 exist *in vivo* in both untreated cells and cells treated with ionizing radiation. We have identified two residues of NBS1, Ser 278 and Ser 343 that are phosphorylated *in vitro* by ATM and whose modification *in vivo* is essential for the cellular response to DNA damage. This response includes S-phase checkpoint activation, formation of the NBS1/Mre11/Rad50 nuclear foci and rescue of hypersensitivity to ionizing radiation. Together, these results demonstrate a biochemical link between cell-cycle checkpoints activated by DNA damage and DNA repair in two genetic diseases with overlapping phenotypes.

The cellular response to DNA damage is complex and includes cell-cycle checkpoint activation, DNA repair and changes in gene transcription⁸⁻¹¹. Cell lines representative of the inherited cancer-prone human diseases ataxia-telangiectasia (A-T) and Nijmegen breakage syndrome (NBS) are hypersensitive to ionizing radiation and have defects in DNA-damage-activated cell-cycle checkpoints¹. Upon DNA damage, ATM phosphorylates p53 (refs 12, 13), and Brca1 (ref. 14). ATM is required for phosphorylation of Chk2 kinase¹⁵ and Rad51 (refs 16, 17) induced by ionizing radiation. NBS1 is an integral component of the Mre11/Rad50/NBS1 nuclease complex⁴⁻⁷ which is important in the repair of DNA double-strand breaks¹⁸.

To examine whether a signalling cascade exists between ATM and NBS1, we studied whether NBS1 was posttranslationally modified following treatment with ionizing radiation. The NBS1 monoclonal antibody specifically recognized a protein with a relative molecular mass of 95,000 (M_r 95K) in all human cell lines examined except those established from NBS patients (data not shown). Although the electrophoretic mobility of the 95K NBS1 protein was constant

throughout the cell cycle, treatment of cells with 10 Gy γ -irradiation resulted in a slower migrating form of NBS1 (Fig. 1a). The mobility of the altered form of NBS1 immunoprecipitated from irradiated cells reverted to that of NBS1 from undamaged cells upon incubation with phosphatase, indicating that NBS1 becomes phosphorylated in response to DNA damage by ionizing radiation (Fig. 1b, compare lane 6 with lanes 2 and 5).

Cell lines lacking functional ATM protein were used to examine whether ATM is required for phosphorylation of NBS1 after DNA damage. The ionizing radiation- or bleomycin (0.1 U per ml)-induced phosphorylation of NBS1 was not detected in these cells

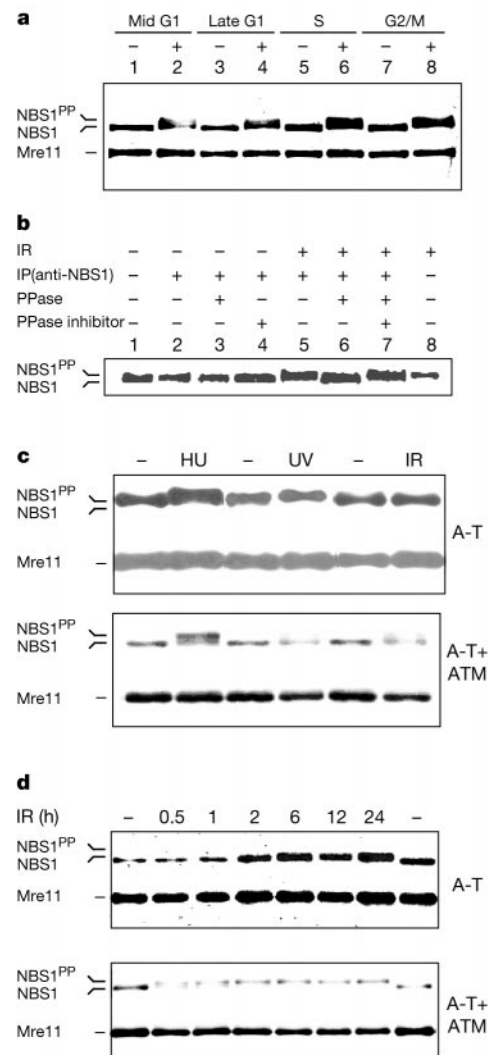


Figure 1 ATM is required for DNA damage-induced phosphorylation of NBS1. **a**, Mobility shift of NBS1 in response to DNA damage during the cell cycle. Synchronized T24 cells were irradiated with 10 Gy γ -irradiation. Lysates (50 μ g protein) of untreated cells and cells 1 h after ionizing radiation (IR) treatment were analysed by western blotting with anti-NBS1 antibody. NBS1^{PP}, phosphorylated form of NBS1. Mre11 is included as a protein loading control. **b**, Phosphorylation and mobility shift of NBS1. Lysates from untreated (lanes 1–4) and IR-treated (lanes 5–8) human lymphoblast NAT10 cells were immunoprecipitated with anti-NBS1 antibody. Immunoprecipitates were incubated with phosphatase (PPase) in the absence or presence of phosphatase inhibitor. Lysates from untreated (lane 1) and IR-treated cells (lane 8) were included as control. **c**, Phosphorylation of NBS1 upon DNA damage and replication block in AT22JIE-T/pEBS7 (A-T) and AT22JIE-T/pEBS7-YZ5 (A-T cells complemented with ATM). **d**, Kinetics of IR-induced modification of NBS1. Cell lysates were prepared at the indicated time points after IR treatment.

[†] Present address: Department of Obstetrics and Gynecology, Kaohsiung Medical University, No. 100, Shih-Chuan 1st Road, Kaohsiung 80708, Taiwan.

(Fig. 1c and data not shown). In contrast, the mobility shift induced by treatment with either hydroxyurea (1 mM) or ultraviolet light (10 J m^{-2}) was normal (Fig. 1c). Importantly, re-introduction of the ATM gene into A-T cells¹⁹ restored phosphorylation of NBS1 induced by ionizing radiation (Fig. 1c, bottom) showing that the failure to phosphorylate NBS1 is directly related to a deficiency of the ATM protein. The absence of ATM does not abolish cellular responses to DNA damage but it does result in a significant delay in these responses²⁰. We therefore compared the kinetics of phosphorylation of NBS1 induced by ionizing radiation in A-T cells, with that in A-T cells that had also been complemented with the ATM protein. Modification of NBS1 was only detectable 6 h after ionizing radiation in the absence of ATM (Fig. 1d). In contrast, the kinetics of NBS1 phosphorylation in A-T cells complemented with ATM was

similar to that of T24 cells, with phosphorylation detectable within 30 min after ionizing radiation (Fig. 1d and data not shown). As the ATM-related kinase, ATR, can phosphorylate p53 with delayed kinetics in the absence of ATM²¹, we suggest that the delayed phosphorylation of NBS1 in A-T cells may be attributable to ATR or another related kinase.

The cellular amounts of the ATM and NBS1 proteins remained unchanged following DNA damage (Fig. 2a, left panel). Although ATM was co-immunoprecipitated by NBS1 antibody in extracts from both treated and untreated cells, ionizing radiation increased the amount of ATM in the NBS1-immunoprecipitates (Fig. 2a, middle). Consistent with these results, NBS1 was co-immunoprecipitated by the ATM antibody from untreated cells and in increased quantities from cells treated with ionizing radiation (Fig. 2a, right).

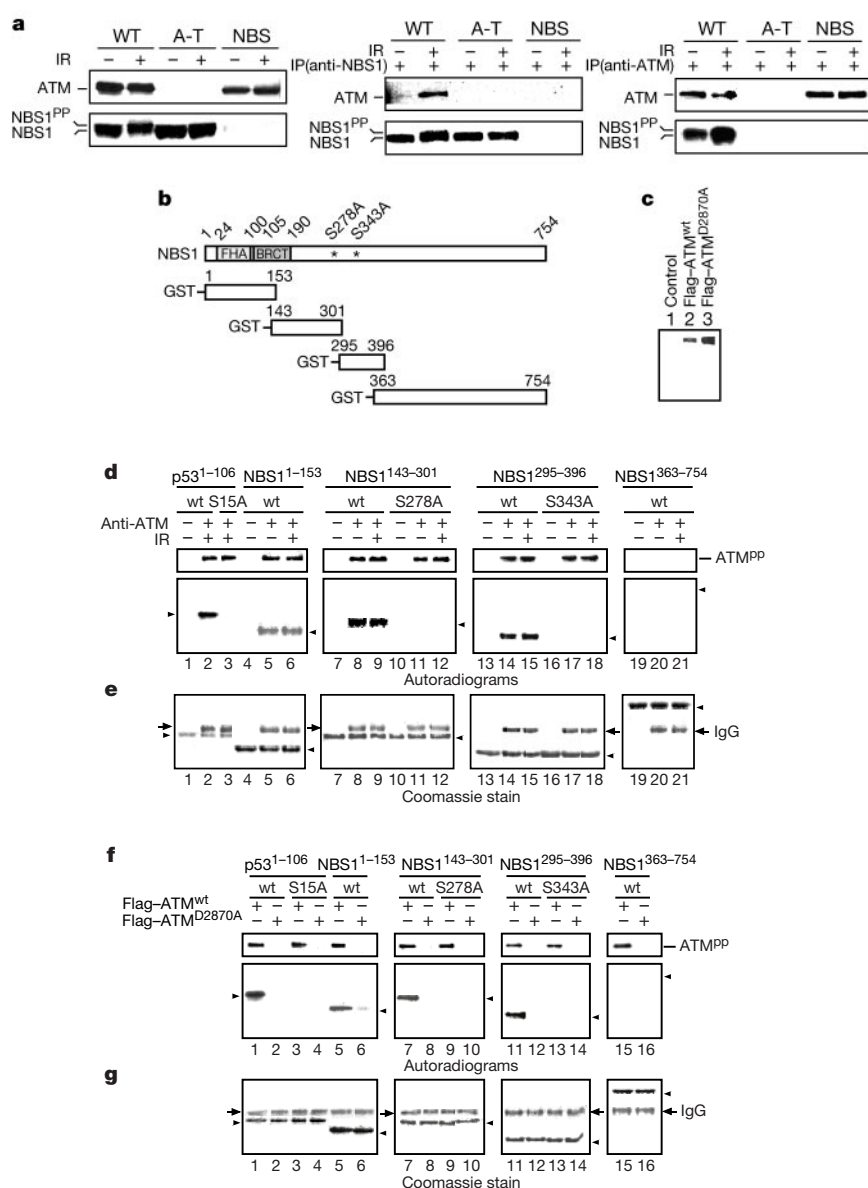


Figure 2 Interaction between ATM and NBS1, *in vivo* and *in vitro* phosphorylation of NBS1 by ATM. **a**, Interaction between ATM and NBS1. Left, direct immunoblotting using cell lysates (100 μg) from mock-treated or 2 h after 10 Gy γ -irradiated cells. T24, A-T (GM09607) and NBS1-LBI cell lines were studied. Middle, co-immunoprecipitation of ATM by anti-NBS1 antibody. Right, co-immunoprecipitation of NBS1 by anti-ATM antibody. **b**, NBS1 structural domains and GST-NBS1 fusion proteins. Asterisk, serine substituted with alanine. **c**, Expression of the recombinant ATM in human 293 cells 24 h after transfection. Immunoblotting with anti-Flag antibody, M5 (Kodak). **d**, Autoradio-

grams showing ATM autophosphorylation (top) and phosphorylation of GST-NBS1 fusion proteins by the ATM immunoprecipitated from either untreated or IR-treated cells (bottom). Purified fusion proteins were incubated with kinase buffer alone (lanes 1, 4, 7, 10, 13, 16 and 19) or kinase buffer and ATM. p53 is a known substrate of ATM. **e**, Coomassie staining showing the amounts of substrates. **f**, Autoradiograms showing Flag-ATM autophosphorylation (top) and phosphorylation of GST-NBS1 fusion proteins by Flag-ATM and Flag-ATM^{D2870A} immunoprecipitates (bottom). **g**, Coomassie staining. Arrowheads, GST fusion proteins; arrows, IgG.

To determine whether ATM can phosphorylate NBS1 *in vitro*, fragments of NBS1 expressed as glutathione *S*-transferase (GST) fusion proteins in *Escherichia coli* were used as substrates in kinase reactions (Fig. 2b) with either endogenous, recombinant Flag-tagged wild-type ATM or a kinase-inactive form of ATM, ATM^{D2870A} (Fig. 2c). Immunoprecipitated ATM was able to phosphorylate itself (Fig. 2d, top) and the GST–NBS1 fusion proteins tested. All phosphorylation occurred within the amino-terminal 396 residues of NBS1 (Fig. 2d, bottom panel).

ATM phosphorylates Ser 15 of p53 (refs 8, 9). Consistent with this result, GST–p53^{1–106}, but not GST–p53^{151A}, was heavily phosphorylated by immunoprecipitated ATM in an *in vitro* kinase assay (Fig. 2d, lanes 2 and 3). Examination of the amino-acid sequences of NBS1 revealed several serine residues followed by glutamines (Q), which are preferred sites for phosphorylation in p53 and c-abl by ATM^{8,9,22}. Alanine substitution of these potential phosphorylation sites within the NBS1 protein identified Ser 278 and Ser 343 as sites phosphorylated by ATM *in vitro* (Fig. 2d, lanes 11, 12, 17 and 18). The site(s) of ATM phosphorylation within amino-acid residues 1–153 of NBS1 is not known but does not appear to be Ser 53 or Ser 58 in the FHA domain, as substitution of these residues with alanine did not change the phosphorylation of GST–NBS1^{1–153} significantly (data not shown).

To provide further evidence that ATM phosphorylates Ser 278 and Ser 343 of NBS1, human 293 cells were transfected with either Flag-tagged wild-type or mutant ATM complementary DNA, and the recombinant proteins were immunoprecipitated with anti-Flag antibody (Fig. 2c, f, top). As expected, wild-type ATM (Fig. 2f, bottom, odd lanes) but not ATM^{D2870A} (Fig. 2f, bottom, even lanes) in the anti-Flag immunoprecipitates could phosphorylate Ser 278 and Ser 343 of GST–NBS1 and Ser 15 of GST–p53.

To determine whether Ser 278 and Ser 343 of NBS1 are phosphorylated *in vivo*, NBS1-LBI, a T-antigen immortalized NBS cell line established from a patient carrying the common founder mutation 657del5 (ref. 23) in NBS1, was transfected with plasmids encoding wild-type NBS1 (pCMV–NBS1^{wt}); NBS1^{S278A} (pCMV–NBS1^{S278A}), NBS1^{S343A} (pCMV–NBS1^{S343A}) or NBS1^{S278A/S343A} (pCMV–NBS1^{S278A/S343A}). As expected, NBS1 protein was not detectable by immunostaining of untransfected cells but was present in the nuclei of cells transfected with either wild-type or mutant

NBS1 constructs (Fig. 3a). Expression of the 95K NBS1 protein was confirmed by immunoblotting of extracts from transfected cells (Fig. 3b). About 20% of the cells were transfected as quantified by immunostaining (data not shown). The protein levels of wild-type and NBS1^{S343A} in the transfected cells are comparable (Fig. 3b, lanes 3–6), but NBS1^{S278A/S343A} expression (Fig. 3b, lanes 9 and 10) was consistently higher than that of the other constructs. Upon treatment with ionizing radiation, the mobility of NBS1^{wt}, but not NBS1^{S278A}, NBS1^{S343A} or NBS1^{S278A/S343A}, was altered (Fig. 3b, compare lanes 5 and 6 with lanes 3 and 4 and 7–10). These data indicate that Ser 278 and Ser 343 are phosphorylated *in vivo*.

Next we examined the effect of NBS1 phosphorylation on cell-cycle checkpoints activated by DNA damage. Both A-T and NBS cells display radioresistant DNA synthesis (RDS)²⁴. RDS is likely to be comparable to the DNA damage- but not replication block-induced S-phase checkpoint. These two distinct checkpoint pathways have been found in various eukaryotes including fission and budding yeast, and mammals^{2–5}. Plasmids encoding either wild-type or mutant NBS1 were co-transfected with an enhanced green fluorescent protein (EGFP) expression plasmid into the NBS1-LBI cells. 25-bromodeoxyuridine (BrdU) was added to cell cultures immediately after mock or ionizing radiation treatment. The ratio of BrdU-positive/EGFP-positive cells to EGFP-positive cells was determined in untreated and treated cells to assess the effect of phosphorylation of NBS1 on RDS. In comparison with wild-type NBS1, NBS1 mutants with one serine residue replaced with alanine were significantly less effective at inhibiting RDS, and an NBS1 mutant with the double alanine substitution was completely ineffective (Fig. 4a).

We examined the effect of NBS1 phosphorylation on cellular sensitivity to ionizing radiation by counting viable cells 72 h after 5 Gy γ -irradiation. Transfection of NBS1-LBI cells with pCMV–NBS1^{wt} resulted in a significant reduction of cellular sensitivity to ionizing radiation compared with the control vector (Fig. 4b). The reduction in cellular sensitivity was abolished by mutation of both the Ser 278 and Ser 343 phosphorylation sites. Interestingly, the sensitivity of the pCMV–NBS1^{S278A} and –NBS1^{S343A} mutants to ionizing radiation was intermediate, indicating that both phosphorylation events contribute to cell sensitivity. Although stable protein complexes of Mre11/Rad50/NBS1 exist in untreated or ionizing

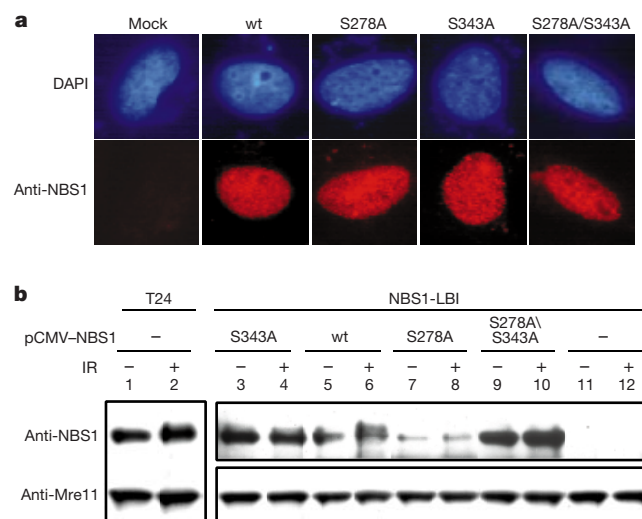


Figure 3 Expression of the full-length wild-type and NBS1 mutant proteins lacking phosphorylation site(s) in NBS1-LBI cells. **a**, Immunostaining of NBS1 proteins in mock or NBS1 expression-plasmid-transfected NBS1-LBI cells. Cells were immunostained with anti-NBS1 monoclonal antibody, MHN1, 36 h after transfection. **b**, Immunoblotting analysis of NBS1 protein in transfected cells. Cells were transfected with expression

plasmids (10 μ g DNA per 2×10^6 cells). Cell treatment is as described in Fig. 2a and cell lysate (100 μ g protein) was analysed. Cell lysates from T24 cells were included as control. Longer exposure is shown to illustrate recombinant NBS1 expression in the transfected NBS1-LBI cells. Shorter exposures are shown for NBS1 and Mre11 in T24 cells, and Mre11 in NBS1-LBI cells.

radiation-treated cells²⁵, it has been reported that Mre11, Rad50, NBS1 and Brca1 all co-localize within nuclear foci induced only after treatment with ionizing radiation^{4,26}. Because NBS1 is required for the formation of foci at DNA damage sites following ionizing radiation^{4,27}, we examined the role of NBS1 phosphorylation in foci formation. Consistent with previous studies²⁵, 26.6% of NBS1-LBI cells transfected with pCMV-NBS1^{WT} were foci-positive (Fig. 4c). In contrast, ionizing radiation did not induce foci formation in cells transfected with pCMV-NBS1^{S278A/S343A} (6.0% foci-positive cells after ionizing radiation, 5.4% foci-positive cells in untreated cells). Similar to the result of the ionizing radiation sensitivity assays, inactivation of either serine phosphorylation site partially restored the ability of NBS1 to form foci in response to ionizing radiation. Our results indicate that phosphorylation of Ser 278 and Ser 343 in NBS1 contributes to both cellular resistance to ionizing radiation and formation of foci that contain NBS1 in response to ionizing radiation.

An enhanced interaction between ATM and NBS1 upon DNA damage was seen using different antibodies in the co-immunoprecipitation experiments. However, we did not detect increased co-

fractionation of ATM and NBS1 upon DNA damage using gel-filtration chromatography (data not shown). It is unclear whether this is due to the transient nature of the enhanced interaction or whether the increase in complexed ATM and NBS1 is the result of conformational changes induced in the protein complex upon DNA damage. We have mapped two residues of NBS1, Ser 278 and Ser 343, which are phosphorylated by ATM. Inactivation of these phosphorylation sites abolishes NBS1 function in DNA damage-activated cell-cycle checkpoints and DNA repair. Phosphorylation of NBS1 by ATM may be required for signal transduction to downstream effectors of DNA damage-induced S-checkpoint activation, such as Chk1 and Chk2 (refs 8–11). We speculate that phosphorylation of NBS1 is also critical for the assembly of repair proteins at the sites of DNA damage. This is supported by the observation that inactivation of the Ser 278 and Ser 343 phosphorylation sites in NBS1 abolished the formation of Mre11/Rad50/NBS1 foci upon treatment with ionizing radiation. The functional interaction between ATM and NBS1 provides a molecular explanation for similar defects in cell-cycle checkpoints and DNA repair exhibited by NBS and A-T cells. Moreover, these studies provide new insights into the complex relationship between DNA damage-activated cell-cycle checkpoints and DNA repair mechanisms that are involved in the cellular response to DNA double-strand breaks.

Methods

Immunoblotting and immunoprecipitation

Cells were lysed in EBC buffer supplemented with protease inhibitors (10 $\mu\text{g ml}^{-1}$ aprotinin, 50 $\mu\text{g ml}^{-1}$ leupeptin, 1 mM phenylmethylsulphonyl fluoride, 100 mM NaF and 1 mM Na_2VO_4). We determined protein concentration by Bradford assay (Biorad). Cell lysates containing 20–100 μg protein were mixed with SDS sample buffer before separation by SDS–7.5% polyacrylamide gel. Proteins were transferred to Immobilon P (Millipore) or nitrocellulose membrane (Schleicher and Schuell). Anti-NBS1 monoclonal antibodies were generated from mice immunized with GST–NBS1^{363–754}. Anti-Mre11 antibodies were generated from mice immunized with purified full-length Mre11 expressed in Sf9 insect cells. For co-immunoprecipitation, we incubated protein lysate (2–3 mg) with primary antibodies overnight at 4 °C as described¹⁷ except that we used secondary antibodies cross-linked to magnetic beads (Dyna).

Phosphatase treatment

Cell lysates containing 1 mg of protein from untreated and ionizing radiation-treated human lymphoblast NAT10 cells were incubated with anti-NBS1 antibody. Immunoprecipitates were washed with EBC buffer and were resuspended in λ phosphatase buffer (New England Biolab) with or without 10 mM phosphatase inhibitor, β -sodium glycerol phosphate.

Kinase assay

ATM was immunoprecipitated from non-irradiated or irradiated HeLa cells and used in kinase assays as described²⁸. GST fusion proteins (2–10 μg) were added to the ATM–protein G sepharose beads and the kinase reaction was carried out in 20 μl reaction volume containing kinase buffer and 10 μCi of $\gamma\text{-}^{32}\text{P}$ -ATP. Recombinant ATM was immunoprecipitated from transfected 293 cells 24 h after transfection using anti-Flag M2 antibody (Kodak).

Plasmid construction

Full-length NBS1 cDNA and the cDNA fragment encoding amino acids 343–754 were generated by polymerase chain reaction with reverse transcription (RT-PCR) using RNA isolated from human Raji cells. The full-length NBS1 cDNA was cloned into a pCMV vector. Site-specific mutagenesis was performed using a QuickChange site-directed mutagenesis kit (Stratagene) and was confirmed by DNA sequencing. To construct the pCMV–Flag–ATM expression vector, a 9.2-kilobase (kb) *Sall*–*XhoI* fragment containing full-length ATM cDNA tagged with a Flag epitope and a His₆ sequence was excised from pFB–YZ5 (ref. 19) and subcloned into a *XhoI*-linearized pCMV vector. Site-specific mutagenesis was performed as described above using a 1.6-kb fragment of the ATM open reading frame as the template. GST–NBS1 and GST–P53 fusion proteins were generated by PCR using full-length cDNAs as templates. The PCR product was cloned into pGEX-4T-3 vector (Pharmacia).

Cell culture and transfection

Primary NBS fibroblasts from Coriell Institute were immortalized with telomerase. The NBS fibroblast cell line NBS1-LBI has been described²². Other cell lines were from the American Type Culture Collection. Cells were cultured in Dulbecco's Modified Eagle medium supplemented with 10% fetal bovine serum (Gibco). Human bladder carcinoma T24 cells were synchronized by density arrest. Cells at various cell cycle stages were

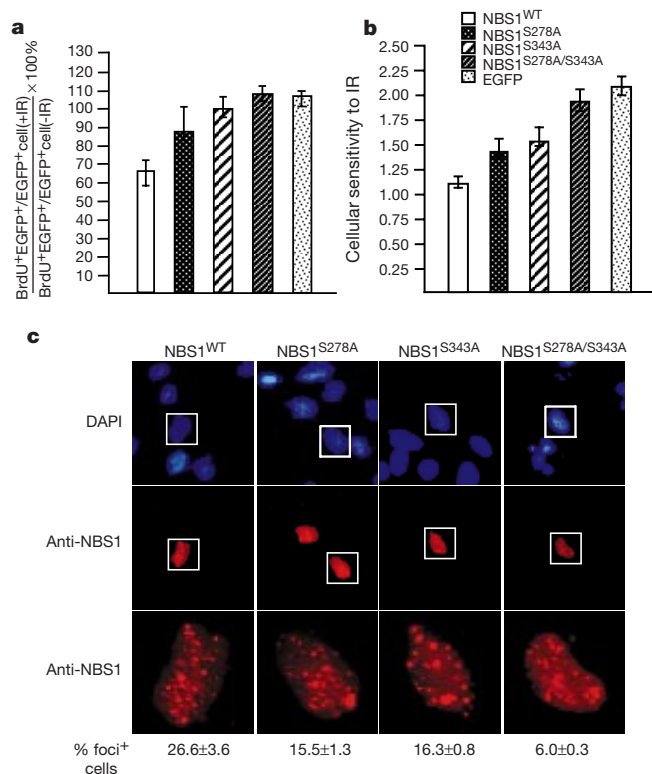


Figure 4 Radioresistant DNA synthesis (RDS) and sensitivity to ionizing radiation by expression of wild-type and NBS1 mutants lacking serine phosphorylation site(s). **a**, RDS. The ratio of BrdU- and EGFP-double positive cells to EGFP-positive cells is determined in untreated and IR-treated cells, respectively. Bar represents the ratio of the IR-treated cells divided by that in the untreated cells multiplied by 100 (>200 fluorescent cells counted per experiment). The mean and s.d. of three experiments is shown. **b**, Cellular sensitivity to IR. Cellular sensitivity to γ -irradiation (5 Gy) was calculated by dividing the number of cells recovered from mock treated cells by the number of cells recovered from the IR-treated cells 72 h after treatment. The mock treated cells have a cellular sensitivity of 1 (>1,000 transfected cells counted per experiment). The mean and s.d. of three experiments is shown. **c**, Formation of IR-induced NBS1 foci in the transfected cells. Three hours after IR (15 Gy), cells were stained with anti-NBS1 monoclonal antibody, MHN1, followed by rhodamine-conjugated anti-mouse antibody. Staining with 4',6'-diamidino-2-phenylindole (DAPI) identifies nuclei (blue). Bottom, $\times 500$ original magnification; top two panels, $\sim \times 31$. Percentages of transfected cells with foci are shown. Similar results were obtained with anti-Mre11 antibody.

collected as described²⁹. Cells were irradiated using a ¹³⁷Cs γ -irradiator at 2.55 Gy per min (Shepherd). Ultraviolet irradiation was performed using UV Stratalinker 2400 (Stratagene). Hydroxyurea (Sigma) was used at 1 mM. Cell lysates were prepared 1 h after treatment with ionizing radiation or ultraviolet unless otherwise specified and after 24 h of hydroxyurea treatment. Transfection was performed using Lipofectamine (Gibco) according to the manufacturer's instructions.

Radioresistant DNA synthesis

NBS1 expression plasmids were co-transfected with pEGFP-N1 (Clontech) in a ratio of 10:1 into the NBS1-LBI cells. Cells were treated with 15 Gy γ -irradiation and immediately incubated in medium containing BrdU for 2 h. We detected BrdU incorporation into DNA using a monoclonal antibody specific for BrdU (Becton-Dickinson) following the immunostaining procedures described below.

Sensitivity to ionizing radiation

Plasmids containing NBS1 cDNA were co-transfected with pEGFP-N1 at a ratio of 10:1 into the NBS1-LBI cells. Cells were treated with 5 Gy γ -irradiation 36 h after transfection. The viable GFP positive cells within 2.9 cm² were counted at 72 h after ionizing radiation.

Immunohistochemistry

Immunostaining was carried out as described¹⁷. Rhodamine-conjugated second antibody was obtained from Jackson Research Laboratories. Slides were mounted in Lipshaw mountant (Immunon) and staining was analysed with a Zeiss AXIOPHOT fluorescence microscope. Images were processed using the software program Adobe Photoshop 5.0.

Received 22 November 1999; accepted 27 March 2000.

- Shiloh, Y. Ataxia-telangiectasia and the Nijmegen breakage syndrome: related disorders but genes apart. *Annu. Rev. Genet.* **31**, 635–662 (1997).
- Gatti, R. A. et al. Localization of an ataxia-telangiectasia gene to chromosome 11q22–23. *Nature* **336**, 577–580 (1988).
- Savitsky, K. et al. The complete sequence of the coding region of the ATM gene reveals similarity to cell cycle regulators in different species. *Hum. Mol. Genet.* **4**, 2025–2032 (1995).
- Carney, J. P. et al. The hMre11/hRad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. *Cell* **93**, 477–486 (1998).
- Varon, R. et al. Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell* **93**, 467–476 (1998).
- Trujillo, K. M., Yuan, S. S., Lee, E. Y. & Sung, P. Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95. *J. Biol. Chem.* **273**, 21447–21450 (1998).
- Paull, T. T. & Gellert, M. Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev.* **13**, 1276–1288 (1999).
- Elledge, S. J. Cell cycle checkpoints: preventing an identity crisis. *Science* **274**, 1664–1672 (1996).
- Weinert, T. DNA damage and checkpoint pathways: molecular anatomy and interactions with repair. *Cell* **94**, 555–558 (1998).
- Caspari, T. & Carr, A. M. DNA structure checkpoint pathways in *Schizosaccharomyces pombe*. *Biochimie* **81**, 173–181 (1999).
- Dasika, G. K. et al. DNA damage-induced cell cycle checkpoints and DNA strand break repair in development and tumorigenesis. *Oncogene* **18**, 7883–7899 (1999).
- Banin, S. et al. Enhanced phosphorylation of p53 by ATM in response to DNA damage. *Science* **281**, 1674–1677 (1998).
- Canman, C. E. et al. Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science* **281**, 1677–1679 (1998).
- Cortez, D., Wang, Y., Qin, J. & Elledge, S. J. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. *Science* **286**, 1162–1166 (1999).
- Matsuoka, S., Huang, M. & Elledge, S. J. Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. *Science* **282**, 1893–1897 (1998).
- Yuan, Z. M. et al. Regulation of Rad51 function by c-Abl in response to DNA damage. *J. Biol. Chem.* **273**, 3799–3802 (1998).
- Chen, G. et al. Radiation-induced assembly of Rad51 and Rad52 recombination complex requires ATM and c-Abl. *J. Biol. Chem.* **274**, 12748–12752 (1999).
- Haber, J. E. The many interfaces of Mre11. *Cell* **95**, 583–586 (1998).
- Ziv, Y. et al. Recombinant ATM protein complements the cellular A-T phenotype. *Oncogene* **15**, 159–167 (1998).
- Canman, C. E., Wolff, A. C., Chen, C. Y., Fornace, A. J. Jr & Kastan, M. B. The p53-dependent G1 cell cycle checkpoint pathway and ataxia-telangiectasia. *Cancer Res.* **54**, 5054–5058 (1994).
- Tibbetts, R. S. et al. A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes Dev.* **13**, 152–157 (1999).
- Baskaran, R. et al. Ataxia telangiectasia mutant protein activates c-Abl tyrosine kinase in response to ionizing radiation. *Nature* **387**, 516–519 (1997).
- Kraakman-van der Zwet, M. et al. Immortalization and characterization of Nijmegen Breakage syndrome fibroblasts. *Mutat. Res.* **434**, 17–27 (1999).
- Young, B. R. & Painter, R. B. Radioresistant DNA synthesis and human genetic diseases. *Hum. Genet.* **82**, 113–117 (1989).
- Maser, R. S., Monsen, K. J., Nelms, B. E. & Petrini, J. H. hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. *Mol. Cell. Biol.* **17**, 6087–6096 (1997).
- Zhong, Q. et al. Association of BRCA1 with the hRad50–hMre11–p95 complex and the DNA damage response. *Science* **285**, 747–750 (1999).
- Nelms, B. E., Maser, R. S., MacKay, J. F., Lagally, M. G. & Petrini, J. H. *In situ* visualization of DNA double-strand break repair in human fibroblasts. *Science* **280**, 590–592 (1998).
- Ziv, Y., Banin, S., Lim, D.-S., Kastan, M. B. & Shiloh, Y. *In Expression and Assay of Recombinant ATM Kinase* (ed. Keyse, S.) (Humana, New Jersey, 1999).

29. Chen, P. L., Scully, P., Shew, J. Y., Wang, J. Y. & Lee, W. H. Phosphorylation of the retinoblastoma gene product is modulated during the cell cycle and cellular differentiation. *Cell* **58**, 1193–1198 (1989).

Acknowledgements

We thank W.-H. Lee, A. Tomkinson and members of their laboratories for discussion; A. Tomkinson, T. Boyer and P. Sung for critical reading; S.-Y. Lee for assistance on the manuscript; M. Chen for site-specific mutagenesis of ATM; Q. Du for the initial studies of ATM kinase; L. Zheng for technical advice; and S. Deb for the full-length human p53 cDNA. E.L. is supported by grants from NIH NINDS, Texas Advanced Research/Advanced Technology Program and NCI P01. S.Z. is supported by a DOD training grant.

Correspondence and requests for materials should be addressed to E.L. (e-mail: leee@uthscsa.edu).

ATM phosphorylation of Nijmegen breakage syndrome protein is required in a DNA damage response

Xiaohua Wu^{*†}, Velvizhi Ranganathan^{‡§}, David S. Weisman^{‡§}, Walter F. Heine^{‡§}, David N. Ciccone^{‡§}, Ted B. O'Neill^{‡||}, Kindra E. Crick^{*†}, Kerry A. Pierce[‡], William S. Lane[‡], Gary Rathbun^{‡||}, David M. Livingston^{*†} & David T. Weaver^{‡§}

^{*} Dana Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA

[‡] Center for Blood Research, 200 Longwood Avenue, Boston, Massachusetts 02115, USA

[†] Departments of Genetics and Medicine, [§] Department of Microbiology and Molecular Genetics and ^{||} Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

[¶] Harvard Microchemistry Facility, Harvard University, Cambridge, Massachusetts 02138, USA

Nijmegen breakage syndrome (NBS) is characterized by extreme radiation sensitivity, chromosomal instability and cancer¹. The phenotypes are similar to those of ataxia telangiectasia mutated (ATM) disease, where there is a deficiency in a protein kinase that is activated by DNA damage, indicating that the Nbs and Atm proteins may participate in common pathways. Here we report that Nbs is specifically phosphorylated in response to γ -radiation, ultraviolet light and exposure to hydroxyurea. Phosphorylation of Nbs mediated by γ -radiation, but not that induced by hydroxyurea or ultraviolet light, was markedly reduced in ATM cells. *In vivo*, Nbs was phosphorylated on many serine residues, of which S343, S397 and S615 were phosphorylated by Atm *in vitro*. At least two of these sites were underphosphorylated in ATM cells. Inactivation of these serines by mutation partially abrogated Atm-dependent phosphorylation. Reconstituting NBS cells with a mutant form of Nbs that cannot be phosphorylated at selected, ATM-dependent serine residues led to a specific reduction in clonogenic survival after γ -radiation. Thus, phosphorylation of Nbs by Atm is critical for certain responses of human cells to DNA damage.

The Atm protein kinase is activated in response to genotoxic stresses. Its substrates include many checkpoint-determining and regulatory proteins, for example, p53, Chk2 and BRCA1 (refs 2–7), indicating its central position in DNA-damage-activated signalling pathways. We investigated whether Nbs acts in the same pathway(s) as Atm and, if so, whether it operates downstream of the latter. Human cells from many tissues, but not NBS cells, synthesize intact Nbs (Fig. 1a). Following γ -radiation, slower SDS–polyacrylamide gel electrophoresis (SDS–PAGE) mobility was observed for Nbs in several tumour-derived and primary fibroblast cell lines (Fig. 1b,

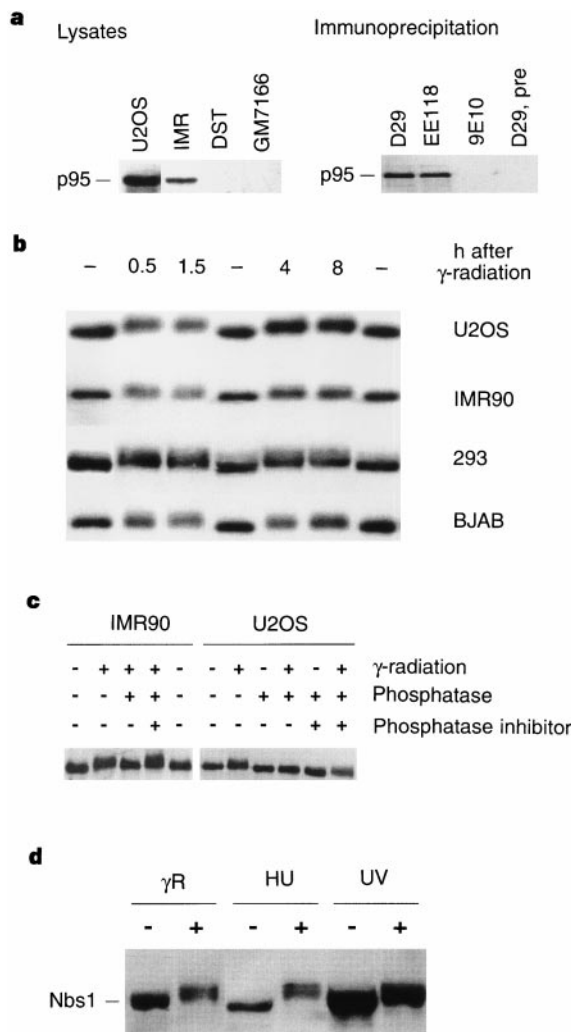


Figure 1 DNA-damage-induced phosphorylation of Nbs. **a**, Lysates from control (U2OS, IMR90) and *NBS* (DST, GM07166) cells were compared by immunoblotting with EE15 (anti-Nbs monoclonal antibody)¹³. U2OS lysates were also immunoprecipitated with D29 (anti-Nbs), EE118 (anti-Nbs), 9E10 (anti-myc) or rabbit pre-immune serum before EE15 immunoblotting. **b**, U2OS, IMR90, 293 and BJAB cells were γ -irradiated (2,000 rad), incubated for 0.5–8 h, and lysates fractionated in 8% SDS–PAGE before EE15 immunoblotting. **c**, Lysates from γ -irradiated IMR90 and U2OS cultures were immunoprecipitated with EE118, and immune complexes incubated with lambda phosphatase (ptase), with or without ptase inhibitor, and detected as above. **d**, 293T cells were treated with γ -irradiation (γ R, 2,000 rad, 1 h), hydroxyurea (HU, 1 mM, 24 h) or ultraviolet light (UV, 50 J m⁻², 1 h) before Nbs detection.

and WS1, HTM1, HBL100, MCF7, Jurkatt, A549 and 293T (data not shown)). The slowest migrating forms were noted at the earliest times measured (30 min after irradiation). *In vitro* incubation with lambda phosphatase led to increased mobility of Nbs; this effect was blocked by co-incubation with phosphatase inhibitors (Fig. 1c). Immunoblotting revealed no phosphotyrosine in Nbs (not shown). Therefore, post-irradiation Nbs phosphorylation probably occurs on serine or threonine residues, or both.

We next tested whether other genotoxic agents that stimulate repair or response pathways, different from those induced by γ -radiation, also induce phosphorylation of Nbs. Ultraviolet light exposure or hydroxyurea treatment (which creates a state of DNA-replication arrest) also resulted in Nbs hyperphosphorylation (Fig. 1d). Like those of γ -radiation, these effects were not restricted to specific cell types (not shown). These findings support the hypothesis that Nbs phosphorylation can be signalled by diverse DNA-damage-induced inputs or from stalled DNA replication forks, and indicate that there may be several upstream pathways that can trigger phosphorylation of Nbs.

Atm, Atr and the DNA-dependent protein kinase (DNAPK) are three related serine/threonine kinases that can be considered as candidate Nbs-phosphorylating enzymes⁸. The protein kinase activities of Atm and DNAPK, but not Atr, are highly sensitive to wortmannin⁹. Incubation of A549 lung carcinoma cells with wortmannin blocked the accumulation of hyperphosphorylated Nbs after irradiation (Fig. 2a) and eliminated the Atm-dependent phosphorylation of Chk2 (not shown). We also analysed *ATM* and *DNAPK* cell lines for DNA-damage-induced phosphorylation of Nbs. *ATM* primary fibroblasts failed to hyperphosphorylate Nbs after irradiation, unlike control fibroblasts and *DNAPK*-deficient glioblastoma cells (Fig. 2b, and data not shown). Moreover, immortalized *ATM* B cells or fibroblasts failed to hyperphosphorylate Nbs, whereas a wild-type *ATM*-reconstituted derivative of one of these cultures¹⁰ re-established Nbs phosphorylating activity following irradiation (Fig. 2b). Treatment with ultraviolet light or hydroxyurea led to hyperphosphorylation of Nbs in *ATM* or control cells (Fig. 2c). Therefore, DNA-damage-induced Nbs hyperphosphorylation requires the Atm kinase for only certain DNA-damage-dependent signalling pathways.

Atm phosphorylates proteins at SQ sites^{3,4,11}. On this basis, Nbs contains several serine residues that are candidate Atm sites including S58, S343, S397 and S615. To define the modifications, we isolated intact Nbs from cells before and after exposure to γ -radiation and identified phosphorylated residues using ion trap mass spectrometry (ITMS). Phosphoserines were unambiguously observed at S343, S397, S432, S509 and S604 in the naive state and after irradiation (Table 1, and see Supplementary Information). We also observed a single phosphorylation at S611, S612 or S615, all of which are present in a single peptide.

Table 1 Mass spectrometry analysis of Nbs phosphorylation

Peptide no.	Peptide	Observed number of PO ₃ groups per peptide	Phosphorylated residue	Ratio <i>ATM</i> /wt
1	TTTPGPSLP S QGVSVDEK	1	S343	< .13
2	MLP S QDAPTVK	1	S397	< .02
3	IPNYQLP S PTK	1	S432	ND
4	IPNYQLSP T KLPSINK	1	(S432, T434)	
5	NKEQHL P SENEPVD T NSDNNLFDTDLK	1	S509	ND
		2	S509, (T516 or S518)	
6	MDIETNDTFSDEAVPESSK S QENEIGKKR	1	(S611, 612, 615)	UD
7	MDIETNDTF P SDEAVPESSK	1	S604	ND

Nbs from naive wild-type BJAB cells was cleaved with trypsin followed by ITMS. Unambiguous phosphopeptides and the stoichiometry of PO₃ groups are shown (phosphorylated residues are in bold and preceded by p) (see Supplementary Information). The same sites were detected after γ -irradiation. Where assignment to particular serine or threonine residues could not be distinguished, residues are underlined. In a second experimental set targeted at S343 and S397 phosphorylated peptides, the ratio of phosphorylation between naive *ATM* (FT169) and wild-type (BJAB) cells was determined. As discussed, it was not feasible to generate equivalent data for peptide 6. ND, not determined; UD, undeterminable.

We tested the ability of Atm to phosphorylate these Nbs phosphorylation sites *in vitro*. Fusions of defined Nbs segments to glutathione *S*-transferase (GST) were generated and the products were purified from bacterial lysates. Active Atm was isolated by immunoprecipitation from cell lysates and served as the enzyme in these assays. GST–Nbs KFRMLS397QDAPT^{392–402} and GST–Nbs TPGPGGLS343QGV^{337–346} were overtly phosphorylated by Atm, and changing a serine to an alanine (S397A or S343A) abrogated the reaction (Fig. 3a). Moreover, a D399A modification flanking S397 decreased the S397–PO₄ signal from this reaction approximately fivefold, in keeping with the Atm selectivity of substrate epitopes¹² (Fig. 3a). S432 (LSP) and S509 (LSE) are not canonical Atm phosphorylation sites^{11,12} and, not surprisingly, GST–NbsKEQHL S509ENEPV^{504–514} was not an Atm substrate *in vitro* (Fig. 3a). GST–Nbs^{495–636}, which contains S611, S612 and S615, was phosphorylated by Atm, and the same protein carrying a mutated S615 residue (S615A) was not efficiently phosphorylated, indicating that there may be an Atm-type phosphorylation site in this peptide (Fig. 3a). S604, S611 and S612 are not canonical Atm sites¹¹ and, given the data of Fig. 3a, the Atm phosphorylation at S615 probably constitutes most of the signal observed using GST–Nbs^{493–636} as substrate.

To assess whether specific sites phosphorylated *in vitro* are also Atm targets *in vivo*, we compared phosphorylation of Nbs S343, S397 and S615 in wild-type and *ATM* cells by ITMS. Although phosphoserine was detectable in each relevant peptide by targeted ion MS/MS (see Methods), there was much less phosphorylation in *Atm* than in wild-type cells at S343 and S397 (about 8–50-fold,

respectively) (Table 1). We could not quantify the relative Atm versus wild-type phosphorylation of S615, at least in part because there is a lysine (K613) between S612 and S615. When S611, S612 or S615 is phosphorylated, our data indicate that this lysine cannot be efficiently cleaved by trypsin, making difficult the comparative MS analysis of this segment from cells that can and cannot efficiently phosphorylate either of these serines. In addition, metabolic labelling with ³²P-orthophosphate indicates that Nbs was phosphorylated about eightfold more in wild-type C3ABR than in the *ATM*-deficient L6 cells (not shown). Collectively, the results strongly support the hypothesis that there are multiple, Atm-preferred phosphorylation sites on Nbs *in vivo*.

To determine whether Atm-dependent phosphorylation sites are functionally important, we expressed full-length wild-type and S→A substitution mutants in NIH3T3 cells^{13,14}. Wild-type human Nbs revealed γ -radiation-dependent phosphorylation and slowing of gel migration; less prominent slowing after γ -radiation was observed with each of the S397A, S615A and S343A proteins (Fig. 3b). As a control, S509A displayed little or no reduction in electrophoretic mobility (Fig. 3b). A triple mutant (S397A+S615A+S509A) showed the same pattern following γ -radiation as S397A or S615A (Fig. 3b). Furthermore, alanine substitutions at S343, S397 or S615 did not diminish the ability of Nbs to associate with Mre11 (Fig. 3c). These findings indicate that at least three sites (S343, S397 and S615) may be responsible for the gel migration change that follows γ -radiation and, together, define the principal Atm-dependent phosphorylations of Nbs.

We next investigated whether cell survival following γ -radiation

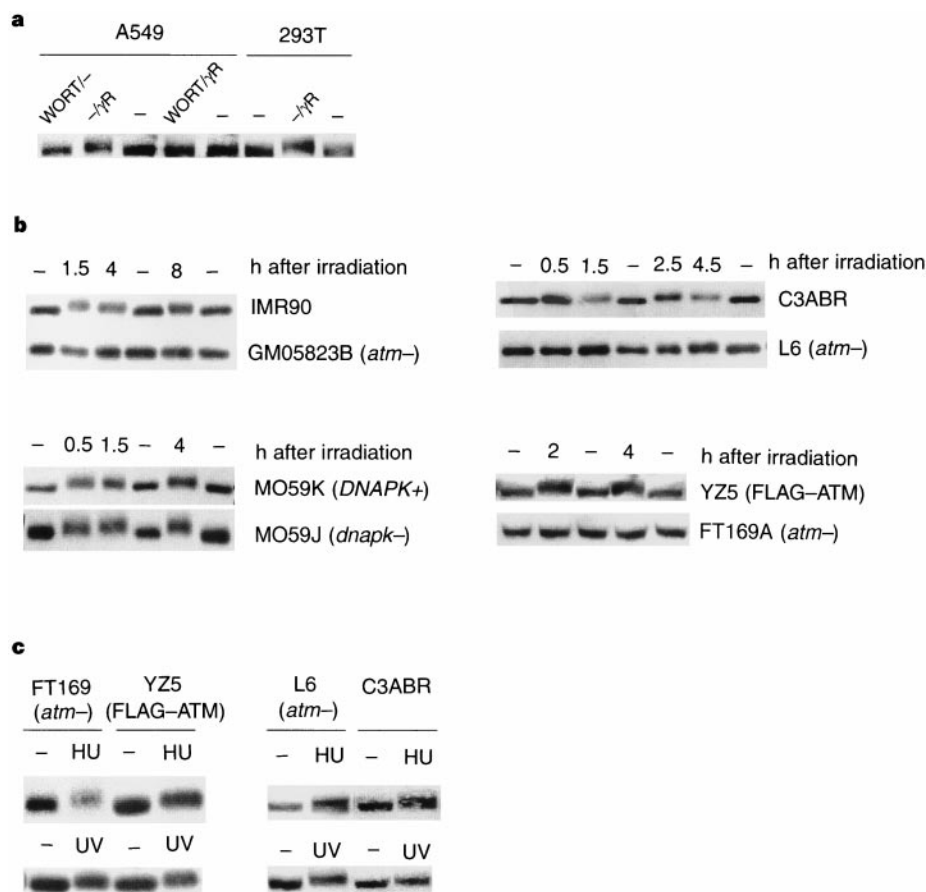


Figure 2 The Atm kinase regulates phosphorylation of Nbs after γ -irradiation. **a**, A549 cells were pre-treated with 30 μ M wortmannin for 30 min and γ -irradiated (γ R, 2,000 rad); lysates were prepared after 30 min and Nbs migration analysed as in Fig. 1. **b**, Nbs from *ATM*, *DNAPK* and control cell lines were compared 0.5–8 h after γ -irradiation (2,000 rad) for IMR90 and GM05823B (control and *ATM* fibroblasts, respectively), MO59K

and MO59J (control and *DNAPK* glioblastoma cells, respectively), C3ABR and L6 (control and *ATM* EBV-immortalized B cells, respectively), FT169A (*ATM*) and YZ5 (FT169A plus FLAG-tagged ATM) fibroblasts. **c**, Nbs migration for C3ABR, L6, FT169A and YZ5 cells following ultraviolet light (UV) or hydroxyurea (HU) treatment.

depends upon phosphorylation of Nbs by Atm by introducing the recombinant wild-type, S397A, S343A and S615A genes into human NBS cells. These cells expressed roughly equivalent levels of the relevant Nbs proteins (see Supplementary Information). Whereas wild-type Nbs corrected the radiosensitivity of the NBS cells, all three of the mutants failed to do so (Fig. 4). Thus, interruption of Atm-linked phosphorylation of Nbs at any of these sites compromises survival after DNA damage, identifying a specific connection between Atm and Nbs in averting cell death after such damage. Finally, the data are consistent with Nbs being an *in vivo* substrate of Atm, although it is formally possible that there are intermediate kinase(s). A critical enzyme–substrate relationship between these two proteins in DNA damage signalling may underlie the previous observation of non-complementation between NBS and ATM cells for the amelioration of radiation hypersensitivity¹.

Nbs is normally associated with the Mre11/Rad50 complex. Loss-of-function mutations of these proteins also result in radiosensitivity and chromosomal instability^{15–17}. NBS deficiency causes failure of Mre11/Rad50 to localize in subnuclear foci after radiation damage¹⁸, but there is no evidence that ATM-dependent Nbs phosphorylation controls the proper intranuclear movement (not shown) or stability (Fig. 3c) of Nbs/Mre11/Rad50 complexes. Nbs stimulates the nuclease activity of Mre11 *in vitro*¹⁹. Therefore, ATM-dependent phosphorylation of Nbs may direct conformational changes of Nbs/Mre11/Rad50 complexes that affect their intrinsic

properties, such as DNA binding and unwinding, nuclease activity and chromatin remodelling.

There is speculation that Nbs/Mre11/Rad50 operates as a sensor of DNA damage and, in that guise, contributes to Atm activation. Our results indicate that Nbs is a direct and/or indirect target of Atm. Although our data do not show that Nbs is an obligate DNA damage sensor required by Atm for its activation, it remains possible that Nbs operates as both a sensor and an Atm target. If Nbs is a target of Atm, there may be a connection between Atm-directed Nbs phosphorylation and activation of radiation-induced checkpoints. Modification of Nbs, like ATM-dependent activation of Chk2, a known checkpoint gene product, occurs very rapidly after irradiation⁵. ATM-dependent Nbs phosphorylation also occurs in unirradiated cells, implying that the relevant signalling pathway is important in normal cellular homeostasis and maintenance of genome integrity. This finding may explain why NBS and ATM patients develop overt disease in the absence of external DNA insults.

Nbs/Mre11/Rad50 complexes are associated with the breast cancer susceptibility protein, BRCA1 (refs 13,20) which is hyperphosphorylated in S phase²¹. Likewise, Atm binds to and phosphorylates BRCA1 at multiple serine residues *in vivo* and *in vitro*⁷. BRCA1 does not appear to regulate the ability of Nbs complexes to form damage-induced nuclear foci¹³, but may control the participation of Atm in homologous recombination and double-

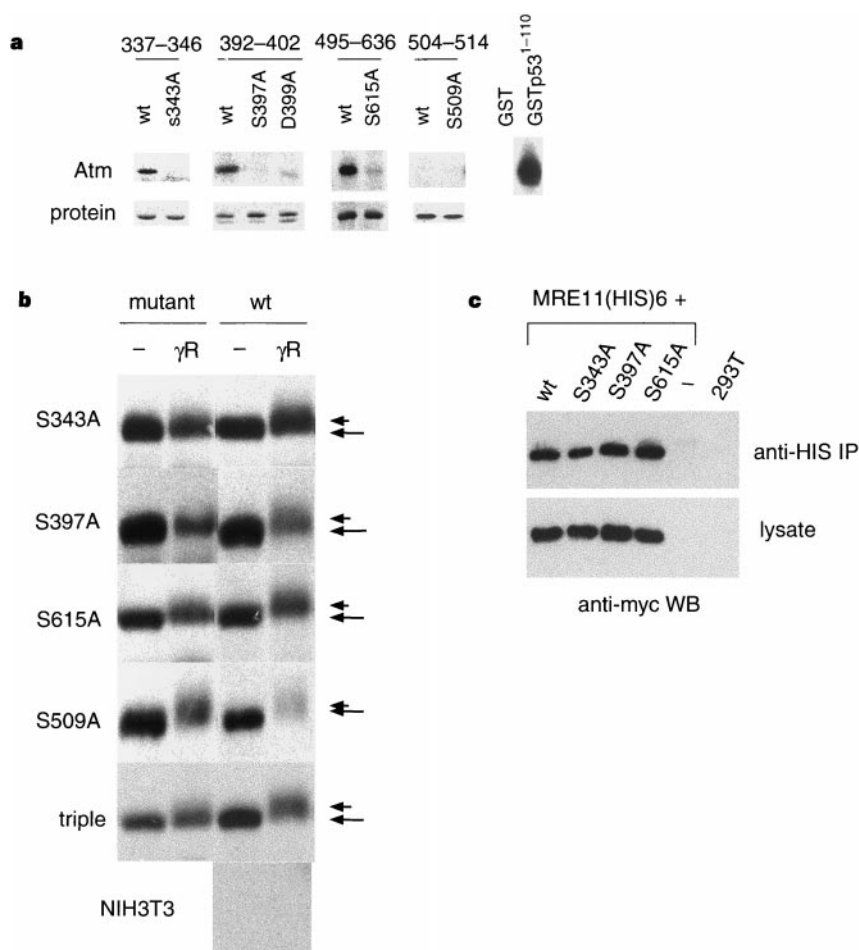


Figure 3 Analysis of possible sites of ATM phosphorylation of Nbs. **a**, *In vitro* Atm kinase activity was tested on GST–Nbs proteins containing the noted Nbs residues, GST–p53^{1–110} or unfused GST, and the input level also displayed ('protein' data). **b**, NIH3T3 cells containing Nbs S397A, S615A, S343A, S509A and an S397A/S509A/S615A triple mutant (triple) were untreated or exposed to γ -radiation, and wild-type

(wt; short arrows) and mutant (long arrows) Nbs detected as in Fig. 1. **c**, 293T cells were co-transfected with expression vectors encoding Mre11–His₆ and myc-epitope-tagged fusions of wild-type, S343A, S397A or S615A Nbs. Lysates were immunoprecipitated with anti-His₆ antibodies and immunoblotted with 9E10 (anti-myc). Expression of transfected Nbs and Mre11 species was constant.

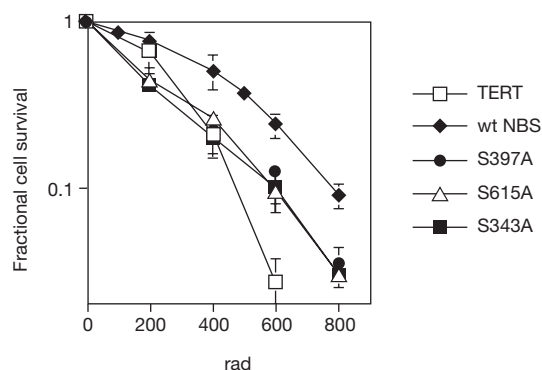


Figure 4 Failure of certain phosphosite mutants of Nbs to protect *NBS* fibroblasts from γ -irradiation-induced death. *NBS* GM07166 TERT and wild-type, S397A, S343A and S615A *NBS* derivative cell lines were treated with 0–800 rad, and surviving colonies counted after three weeks. Fractional cell survival is the fraction of colonies surviving irradiation divided by the total colony number in an unirradiated parallel culture. Standard deviation is derived from three independent experiments. All four alleles were equivalently expressed in these cells, as measured by comparative immunoprecipitation or immunoblotting with Nbs polyclonal or monoclonal antibodies (see Supplementary Information).

stranded break repair^{22–25}, which are also Nbs-dependent functions. These proteins may have shared contributions to tumour prevention, as there is an elevated cancer susceptibility for *BRCA1*, *NBS* and, possibly, *ATM* carriers^{26,27}.

Methods

Cell lines

Control (IMR90, WS1, HTM1), *NBS* (GM07166) and *ATM* (GM05823, GM08391) fibroblasts were obtained from Coriell Human Mutant Cell Repository or ATCC. 293, 293T, U2OS, SAOS-2, M059K, M059J (*DNAPK*–) and GM07166 telomerase catalytic subunit (TERT) cells were cultivated as described^{14,21,22}. A549 lung carcinoma, G361 melanoma, DST (*NBS*), L6 (*ATM*) and C3ABR (control) EBV-B cell lines, and FT169A *ATM* fibroblasts and their derivative, YZ5 (*ATM* cDNA), were provided by R. Abraham, P. Concannon and Y. Shiloh.

Nbs antibodies and Nbs detection

Polyclonal (D29) and monoclonal antibodies (EE15, EE118) were raised against a human Nbs fragment fused to glutathione S-transferase (GST–Nbs^{420–774}). Lysates from 2×10^6 cells were prepared in ice-cold 0.5% NP40, 150 mM NaCl, 50 mM Tris–Cl, pH 8.0, containing protease and phosphatase inhibitors, and electrophoresed in 8% SDS–polyacrylamide gels before immunoblotting. Alternatively, lysates were immunoprecipitated for 2 h at 4°C, followed by 1 h adsorption to Protein A-Sepharose. Protein phosphatase treatment was conducted by incubating D29 immune complexes with 500 U of lambda phosphatase (Boehringer–Mannheim) with or without 25 mM NaF and 25 mM sodium orthovanadate for 2 h at 30°C.

EE118 antibody–Sepharose coupling was achieved after adsorption of hybridoma supernatant to Protein A-Sepharose, washing with 0.1 M sodium borate, pH 9.0, and conjugation for 1 h at room temperature in 40 mM dimethyl pimelimidate dihydrochloride, 0.1 M sodium borate, pH 9.0. The beads were then incubated with 40 mM ethanolamine, 0.1 M sodium borate, pH 8.0 for 2 h at room temperature, washed three times with NETN and stored at 4°C in NETN, 0.1% sodium azide.

Ion trap mass spectrometry

BJAB or *ATM*-deficient FT169A cells (5 ml packed cell volume) were lysed in 0.5% NP40, 170 mM NaCl, 50 mM Tris–Cl, pH 8.0, $1 \mu\text{g ml}^{-1}$ aprotinin, $50 \mu\text{g ml}^{-1}$ PMSF, $1 \mu\text{g ml}^{-1}$ leupeptin, 0.2 mM sodium orthovanadate, 4 mM NaF for 15 min. EE118 beads were used to immunoprecipitate Nbs (3 h, 4°C). BJAB cells were also isolated following γ -radiation (2,000 rad, 4 h). The precipitates were washed three times with lysis buffer, boiled in 1% SDS, 50 mM Tris–Cl pH 7.5, 5 mM DTT, 5 mM EDTA, diluted with 10 volumes of lysis buffer, and a second Nbs immunoprecipitation was then conducted overnight at 4°C. Samples were fractionated by SDS–PAGE and recovered by excising the relevant Coomassie blue-stained bands. Control samples without the coupled antibody did not reveal proteins at 95–100K.

Proteins were subjected to in-gel reduction, carboxyamidomethylation and tryptic digestion. Phosphorylated peptides were detected and sequenced using bimodal chromatography on an Fe^{3+} metal affinity/reverse phase microcolumn coupled to an LCQ Deca quadrupole ion trap mass spectrometer (ThermoQuest), using a custom nanoelectrospray

source (LC/MS/MS). The ion trap repetitively surveyed the range 395–1,295 m/z , executing data-dependent scans on the three most abundant ions in the survey scan, allowing high-resolution (zoom) scans to determine charge state and exact mass, and MS/MS spectra for peptide sequence information. MS/MS spectra were acquired with a relative collision energy of 30% and an isolation width of 2.5 D. Recurring ions were dynamically excluded. After database correlation with SEQUEST, phosphorylated peptides were confirmed by manual, *de novo* interpretation of the MS/MS spectra using FuzzyIons²⁸.

Second sets of LC/MS/MS experiments were conducted for each tryptic peptide containing putative phosphorylation at S343, S397 and S611, S612 and S615 to determine the relative ionization intensity ratio of phosphorylated to unphosphorylated peptides. In each of these experiments, two alternating scan events were acquired. In the first, an MS/MS scan is targeted for a given phosphorylated peptide and the structural information from this scan acts as a chromatographic locator for the peptide. The second full scan event provided the ion intensity ratios for each phosphorylated/unphosphorylated pair.

In vitro *Atm* kinase assays

Nbs protein fragments were converted to GST fusion proteins by cloning PCR products into pGEX4T2. Protein concentrations and purity were estimated by Bradford assay and Coomassie blue staining of SDS–PAGE, respectively. *Atm* kinase was immunoprecipitated from G361 cells using anti-*Atm* antibody (AB3, Oncogene Sci)⁹. *In vitro* kinase assays¹² were conducted on equivalent quantities of target protein.

Recombinant Nbs expression

WZL–NBS (NBS in a WZL retroviral expression vector¹⁴), single residue derivatives (S343A, S397A, S615A and S509A) or a triple mutation (S397A/S509A/S615A) were transfected into 293 cells with VSV G and GAG, POL-expressing plasmids¹⁴. Supernatants were harvested after 48–96 h and used to infect NIH3T3 or GM07166 TERT cells. Infected cells were selected in $1 \mu\text{g ml}^{-1}$ blasticidin (Invitrogen). Telomerase catalytic subunit (TERT) expression allows continued cell proliferation of certain primary human cells¹⁴.

Received 30 December 1999; accepted 19 April 2000.

- Shiloh, Y. Ataxia-telangiectasia and the Nijmegen Breakage Syndrome—related disorders but genes apart. *Annu. Rev. Genet.* **31**, 635–662 (1997).
- Savitsky, K. *et al.* A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* **268**, 1749–1753 (1995).
- Banin, S. *et al.* Enhanced phosphorylation of p53 by ATM in response to DNA damage. *Science* **281**, 1674–1677 (1998).
- Canman, C. E. *et al.* Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science* **281**, 1677–1679 (1998).
- Matsuoka, S., Huang, M. & Elledge, S. J. Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. *Science* **282**, 1893–1897 (1998).
- Brown, A. L. *et al.* A human Cds1-related kinase that functions downstream of ATM protein in the cellular response to DNA damage. *Proc. Natl. Acad. Sci. USA* **96**, 3745–3750 (1999).
- Cortez, D., Wang, Y., Qin, J. & Elledge, S. J. Requirement of ATM-dependent phosphorylation of Brca1 in the DNA damage response to double-strand breaks. *Science* **286**, 1162–1166 (1999).
- Hoekstra, M. F. Responses to DNA damage and regulation of cell cycle checkpoints by the ATM protein kinase family. *Curr. Opin. Genet. Dev.* **7**, 170–175 (1997).
- Sarkaria, J. *et al.* Inhibition of phosphoinositide 3-kinase related kinases by the radiosensitizing agent wortmannin. *Cancer Res.* **58**, 4375–4382 (1998).
- Ziv, Y. *et al.* Recombinant ATM protein complements the cellular A-T phenotype. *Oncogene* **15**, 159–167 (1997).
- Kim, S. -T., Lim, D. -S., Canman, C. E. & Kastan, M. B. Substrate specificities and identification of putative substrates of ATM kinase family members. *J. Biol. Chem.* **274**, 37538–37543 (1999).
- O'Neil, T. *et al.* Utilization of oriented peptide libraries to identify substrate motifs selected by ATM and DNA-PK. *J. Biol. Chem.* (in the press).
- Wu, X., Petrini, J. H. J., Weaver, D. T., Livingston, D. M. & Chen, J. Independence of Nbs1/Mre11/Rad50 nuclear focus formation and the presence of intact BRCA. *Science* (in the press).
- Ciccone, D. N. *et al.* Rescue of a telomere length defect of Nijmegen breakage syndrome cells by NBS and Telomerase catalytic subunit. *Nature Genet.* (submitted).
- Stewart, G. *et al.* The DNA double-strand break repair gene hMRE11 is mutated in individuals with an ataxia-telangiectasia-like disorder. *Cell* **99**, 577–587 (1999).
- Luo, G. *et al.* Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. *Proc. Natl. Acad. Sci. USA* **96**, 7376–7381 (1999).
- Yamaguchi-Iwai, Y. *et al.* Mre11 is essential for the maintenance of chromosomal DNA in vertebrate cells. *EMBO J.* **18**, 6619–6629 (1999).
- Carney, J. P. *et al.* The hMre11/Rad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. *Cell* **93**, 477–486 (1998).
- Paull, T. T. & Gellert, M. Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev.* **13**, 1276–1288 (1999).
- Zhong, Q. *et al.* Association of BRCA1 with the hRad50-hMre11-p95 complex and the DNA damage response. *Science* **285**, 747–750 (1999).
- Scully, R. *et al.* Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90**, 425–435 (1997).
- Scully, R. *et al.* Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**, 265–275 (1997).
- Chen, G. *et al.* Radiation-induced assembly of Rad51 and Rad52 recombination complex requires ATM and c-Abl. *J. Biol. Chem.* **274**, 12748–12752 (1999).
- Moynahan, M. E., Chiu, J. W., Koller, B. H. & Jasin, M. Brca1 controls homology-directed DNA repair. *Mol. Cell* **4**, 511–518 (1999).
- Scully, R. *et al.* Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol. Cell* **4**, 1093–1099 (1999).

26. Seemanova, E. An increased risk for malignant neoplasms in heterozygotes for a syndrome of microcephaly, normal intelligence, growth retardation, remarkable facies, immunodeficiency and chromosomal instability. *Mutat. Res.* **238**, 321–324 (1990).
27. Athma, P., Rappaport, R. & Swift, M. Molecular genotyping shows that ataxia-telangiectasia heterozygotes are predisposed to breast cancer. *Cancer Genet. Cytogenet.* **92**, 130–134 (1996).
28. Chittum, H. S. *et al.* Rabbit b-globin is extended beyond its UGA stop codon by multiple suppressions and translational reading gaps. *Biochemistry* **37**, 10866–10870 (1998).

Supplementary information is available on Nature's World-Wide Web site (<http://nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank P. Concannon, Y. Shiloh and R. Abraham for distribution of cell lines; J. Gan and J. DeCaprio for help with monoclonal antibody production; members of the Livingston and Weaver labs for discussions; and R. DePinho, D. Kufe, F. Alt, S. Kharbanda and M. Weber for comments and criticism and for reading the manuscript. This work was supported by funds from the NIH and the March of Dimes Foundation. X.W. was supported by a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research.

Correspondence and requests for materials should be addressed to D.M.L. or T.W.

Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene

Adam C. Bell & Gary Felsenfeld

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0540, USA

The expression of the insulin-like growth factor 2 (*Igf2*) and *H19* genes is imprinted. Although these neighbouring genes share an enhancer¹, *H19* is expressed only from the maternal allele, and *Igf2* only from the paternally inherited allele^{2,3}. A region of paternal-specific methylation upstream of *H19* appears to be the site of an epigenetic mark that is required for the imprinting of these genes^{4,5}. A deletion within this region results in loss of imprinting of both *H19* and *Igf2* (ref. 5). Here we show that this methylated region contains an element that blocks enhancer activity. The activity of this element is dependent upon the vertebrate enhancer-blocking protein CTCF. Methylation of CpGs within the CTCF-binding sites eliminates binding of CTCF *in vitro*, and deletion of these sites results in loss of enhancer-blocking activity *in vivo*, thereby allowing gene expression. This CTCF-dependent enhancer-blocking element acts as an insulator. We suggest that it controls imprinting of *Igf2*. The activity of this insulator is restricted to the maternal allele by specific DNA methylation of the paternal allele. Our results reveal that DNA methylation can control gene expression by modulating enhancer access to the gene promoter through regulation of an enhancer boundary.

Maternal transmission of a 1.6-kilobase (kb) deletion, within a differentially methylated region which is located between the mouse *Igf2* and *H19* genes, results in expression of the normally silent *Igf2* allele⁵. We examined the ability of this deleted fragment (DMD) to act as a positional enhancer-blocking element by inserting it at various locations relative to an enhancer in our model system^{6–8}, described in Methods. Insertion of the 1.6-kb DMD fragment between the enhancer and the promoter results in an 8–10-fold drop in colony number, similar to the 8-fold drop observed with the previously characterized 1.2-kb chicken β -globin insulator⁷. These results are not attributable to the increased distance between the enhancer and the promoter, as insertion of up to 2.3-kb of heterologous DNA between them has little effect on colony number⁷.

Furthermore, like the β -globin insulator⁸, placing the DMD outside the enhancer–promoter path, either upstream of the enhancer or downstream of the promoter, has little effect on expression (Fig. 1b). From these results, we conclude that the DMD has the position-dependent enhancer-blocking properties of an insulator.

The DMD fragment is part of a larger (~2-kb) imprinted control region (ICR) that is methylated exclusively on the paternal allele throughout development^{4,9}. Allele-specific alterations in chromatin structure also occur in this region^{10–12}. Two nuclease-hypersensitive regions are located on the maternal allele (HS1 and HS2 in Fig. 1a), whereas the chromatin on the paternal allele is methylated and nuclease insensitive¹². Both HS1 and HS2 remain hypersensitive throughout development and are present independent of tissue type. We tested the enhancer-blocking potential of fragments spanning HS1, HS2 and a larger fragment that spans the entire ICR. All of these fragments confer enhancer-blocking activity (Fig. 1b). HS1 and HS2 individually show considerable enhancer-blocking activity, a fragment that contains both HS1 and HS2 essentially eliminates the enhancer's influence on expression (Fig. 1b, compare NIΔE with ICR).

A binding site (FII) for CTCF within the core of the chicken β -globin insulator is an essential component of the enhancer-blocking activity of that element⁶. A comparison between FII and 2.6-kb of mouse sequence spanning the ICR revealed a 13/16

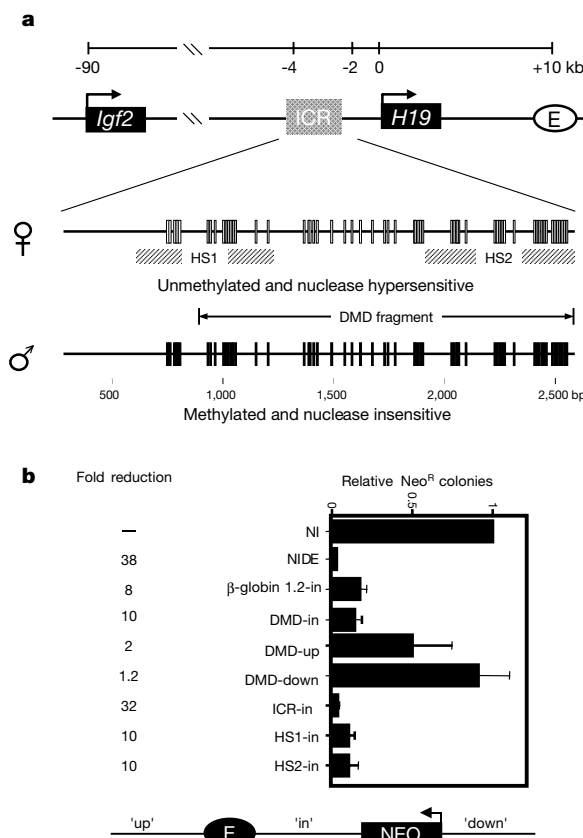


Figure 1 A differentially methylated region upstream of *H19* has the enhancer-blocking properties of an insulator. **a**, Neighbouring mouse *Igf2* and *H19* genes. On the maternally inherited chromosome, the ICR is unmethylated (white rectangles) and contains two nuclease-hypersensitive regions (hatched boxes, HS1 and HS2); on the paternally inherited chromosome, the ICR is methylated (black rectangles) and contains no hypersensitive sites¹². Deletion of a 1.6-kb fragment of the ICR (the DMD fragment) eliminates HS2 and most of HS1 (ref. 5). **b**, Enhancer-blocking activity was measured by a standard assay^{6–8} in which fragments of the ICR were inserted at defined positions relative to the enhancer and promoter in stably integrated constructions (see Methods). For each construction, colony number was normalized to un-insulated control, NI. Data are the average of three independent measurements.

base-pair (bp) match between the 3' end of FII and a sequence at the 5' edge of HS2 (m3 in Fig. 2a). By searching the remainder of the ICR with the m3 sequence, we identified four homologous sequences (m1–4 in Fig. 2a). Sequences homologous to these mouse sites are also found upstream of the human and rat *H19* genes, indicating that they are functionally important (aligned in Fig. 2a). The conservation of sequences overlapping those shown here has been noted^{13,14}. In humans, a region of paternal-specific DNA methylation upstream of *H19* has also been described^{15,16} and the homologous sequences defined here are part of a larger repeating element found in that region. An alignment of all of the sites from rat, human and mouse reveals a 12-bp consensus sequence that is shared among them (Fig. 2a). This consensus sequence bears a 11/12-bp match with the sequence at the 3' end of β -globin FII. Notably, this 3' region of FII makes an essential contribution to CTCF binding and enhancer blocking⁶. Consistent with these observations, we find that each of these mouse sites and a representative human site bind to both purified chicken CTCF (P in Fig. 2c) and *in vitro* translated human CTCF (I in Fig. 2c). Like FII, a fragment spanning a single mouse ICR site confers position-dependent enhancer-blocking activity *in vivo* (Fig. 2b). Further-

more, in gel shifts with K562 nuclear extracts (E in Fig. 2c) (the cells in which the enhancer-blocking assays were performed), a complex that co-migrates with the FII/CTCF complex was observed with each of these mouse and human ICR sites (Fig. 2c), and an antibody raised against CTCF supershifted this complex (Fig. 2d). Consistent with our attribution of enhancer-blocking activity to CTCF, this complex was competed by FII, but not by a mutant of FII in which both enhancer blocking and CTCF binding have been eliminated (FIIx3' in Fig. 2d)⁶. Alteration of the base pairs shared between the ICR sites, and FII, in the context of one of these mouse sites, resulted in loss of competition for binding to CTCF (m1x3' in Fig. 2d). Mutation of the same base pairs in the context of a human ICR site has the same effect on CTCF binding¹⁷.

In the mouse *H19* ICR, HS1 and HS2 each contain two CTCF sites. We sequentially deleted each of these sites (Fig. 3), and measured the enhancer-blocking activities of the resulting fragments. As the enhancer-blocking activities of HS1 and HS2 are dependent upon their orientations (data not shown), deletion analyses were carried out with the orientation that gave the strongest activities. In each case, a deletion that eliminates either of the CTCF sites results in a reduction in enhancer-blocking

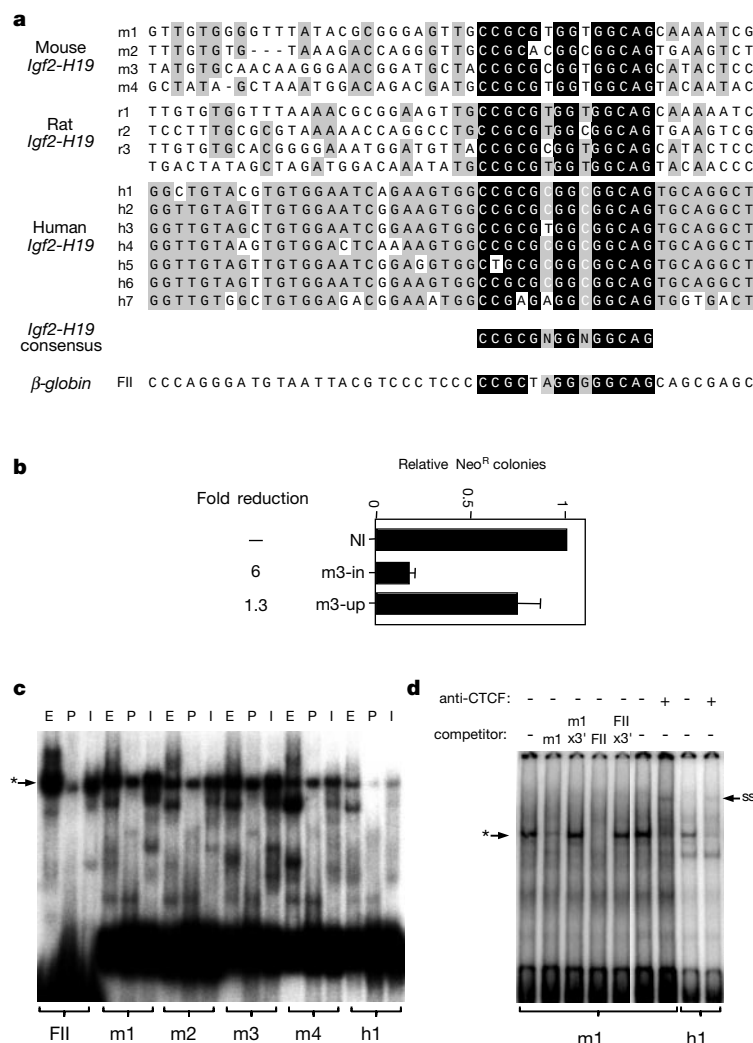


Figure 2 Conserved CTCF sites within the *H19* ICR. **a**, Sequences of the CTCF sites clustered upstream of the mouse, rat and human *H19* genes. Shading indicates identity among the sites: grey, species-specific identities; black, cross-species sequence conservation among these sites. **b**, Enhancer-blocking activity of a fragment spanning only m3 from the mouse ICR. Data are the average of three independent experiments. **c**, Gel mobility-shift analysis of β -globin FII (60-mer) and 83–91-mer duplexes spanning

mouse and human ICR sites binding to K562 nuclear extract (E), partially purified (chicken) CTCF (P) and *in vitro* translated chicken CTCF (I). Asterisk, the position of the CTCF–DNA complex. Labeled DNA probes are indicated at the bottom. **d**, Analysis of CTCF binding to representative mouse and human ICR sites. DNAs were incubated with K562 nuclear extract in the presence of a 50-fold excess of unlabelled competitors as indicated or with an anti-CTCF antibody. SS, position of the supershifted CTCF complex.

activity, whereas deletion of both sites from either HS1 or HS2 eliminates their activity (Fig. 3a). The deletions that we made span sequences that are larger than the average 53-bp CTCF footprint. Among these sequences, however, the only significant similarity is within the CTCF sites. These similarities define a consensus sequence for CTCF binding (Fig. 2a, d), and we have shown that this consensus sequence is essential for the enhancer-blocking activity of β -globin FII⁶. We have also shown that single CTCF sites from several other loci (including the mouse ICR, Fig. 2b) alone confer enhancer-blocking activity in our assay⁶. Consistent with these findings, a human ICR site also confers enhancer-blocking activity and point mutations that eliminate the CTCF consensus from this site result in loss of enhancer-blocking activity *in vivo*¹⁷.

We have shown that sequences within the mouse H19 ICR have the enhancer-blocking properties of an insulator. Several observations suggest that this activity is directly involved in the regulation of *Igf2*. If the *H19* enhancer is moved from its genomic location downstream of *H19* to a new location upstream of the ICR, the normally silent maternal allele of *Igf2* is expressed¹⁸. This indicates that the enhancer's position downstream of the *H19* locus may prevent activation of the maternal *Igf2* allele. Competition between the *H19* and *Igf2* promoters cannot explain this result because deletion of the *H19* promoter has no effect on *Igf2* expression¹⁹. Instead, the enhancer's position relative to the ICR restricts its action¹⁸: a deletion within the ICR results in biallelic expression of *Igf2* (ref. 5). This proposal is further supported by the observation that maternal inheritance of the relocated enhancer results in loss of expression of the normally active *H19* allele. In this case, because the ICR is now located between the enhancer and the *H19* promoter, it

blocks their interaction. Thus, the dependence of *H19* and *Igf2* expression on the position of the *H19* enhancer is explained by a model that posits the existence of an insulator within the ICR^{5,18,20}.

The influence of the ICR on expression of *Igf2* depends upon the allele's parent of origin. It has been proposed that this is because the *H19* locus contains an insulator that is active only on the unmethylated (maternal) allele²⁰. In this model, inheritance of paternal-specific CpG methylation in the ICR results in inactivation of the insulator and thus the *H19* enhancer is free to activate *Igf2* on this allele. Direct support for a role of DNA methylation in activation of *Igf2* comes from the observation that in DNA methyltransferase-1-deficient mouse embryos, both alleles of *Igf2* are silent²¹.

The results of our deletion analysis of HS1 and HS2 indicate that the conserved CTCF sites in these elements may be responsible for their enhancer-blocking activities. One model that could explain why CpG methylation abolishes this activity is that CTCF cannot bind these sites when they are methylated. To test this, we made the corresponding oligomers with 5-methyl cytosine incorporated at each CpG, and assessed the ability of the resulting duplex to compete for the binding of CTCF to the unmethylated form. Methylation of each of the mouse sites, and a representative human site, greatly reduces their ability to compete for binding of CTCF to an unmethylated site, even at a 50-fold molar excess (Fig. 3b). Methylation of β -globin FII has a similar effect. Because FII and the ICR sites have only one CpG in common, we examined the influence of methylation at only this site (on both strands) in several ICR sites (Fig. 3b, right). In fact, methylation of this CpG alone significantly reduces CTCF binding to all of these sites (Fig. 3b, M1 lanes). This result indicates that enhancer access could, in principle, be regulated by a single (perhaps targeted) methylation event. Further examination of the influence of methylation on gene expression will require a system that allows the establishment and maintenance of partially methylated transgenes *in vivo*.

We have shown that the *H19* ICR is an enhancer-blocking element. CTCF-binding sites are required for this activity and when these sites are methylated they no longer bind the insulator protein CTCF. Our results, and those of ref. 17, provide direct evidence for a mechanism for *Igf2* imprinting, in which differential methylation of an enhancer boundary allows epigenetic control of *Igf2* expression in the embryo (Fig. 4). In humans, a causal link between overexpression of *Igf2* and the pathogenesis of some cases of Beckwith–Wiedemann syndrome has been suggested^{22–25}. Beckwith–Wiedemann syndrome, or fetal overgrowth syndrome, is a disorder of prenatal overgrowth and predisposition to embryonic malignancies such as Wilms tumour. There is a correlation between loss of imprinting of *Igf2* in Wilms tumour and Beckwith–Wiedemann

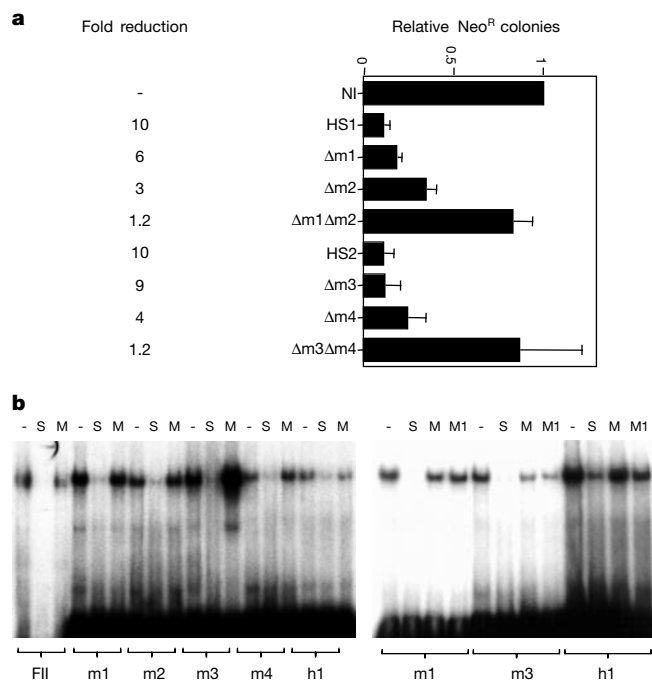


Figure 3 CTCF is responsible for the methylation-sensitive enhancer-blocking activities of the mouse and human ICRs. **a**, Enhancer-blocking activities of fragments of the mouse ICR after sequential deletion of the sequences spanning individual CTCF sites. Results are the average of two to three independent measurements. **b**, Left, effect of CpG methylation on binding of partially purified chicken CTCF to various sites in the absence of competitor DNA (-), or in the presence of a 50-fold excess of unlabelled duplex DNA (S, self) or a 50-fold excess of unlabelled duplex of identical sequence with 5-methyl C incorporated at every CpG (M, uniformly methylated) (identical results were obtained with human CTCF (data not shown)). Labelled DNA probes are indicated at the bottom of the panel. Right, the effect of 5-methyl C substitution at a single site (M1, singly methylated at the first CpG) in the black-shaded region of Fig. 2a).

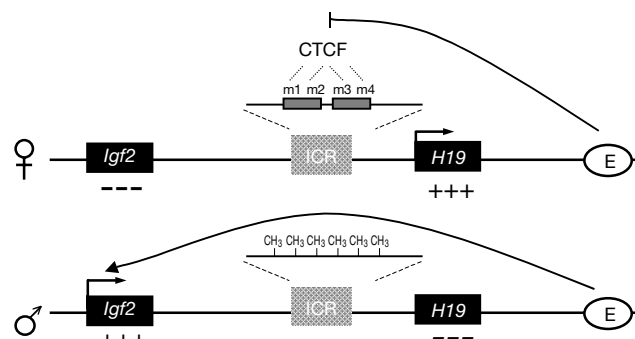


Figure 4 A model for methylation-dependent modulation of insulator action in the epigenetic regulation of *Igf2*. On the maternally inherited chromosome, the ICR is unmethylated. This allows binding of CTCF to its sites (m1–4), two in each nuclease-hypersensitive region (shaded boxes), and the resulting insulator blocks activation of the maternal copy of *Igf2* by the *H19* enhancer. On the paternally inherited chromosome, the ICR is methylated. This prevents CTCF binding, thereby inactivating the insulator and allowing the *H19* enhancer to activate *Igf2*.

syndrome and increased methylation of the maternal *H19* allele^{26–29}. In Wilms tumour, this aberrant methylation pattern includes the CTCF sites illustrated in Fig. 2a¹⁶. These sites are consistently methylated on both alleles in Wilms tumours with loss of *Igf2* imprinting. Our data are consistent with the idea that the loss of *Igf2* imprinting observed in those tumours is caused by inactivation of a CTCF-dependent insulator in that locus.

In *Drosophila*, the activity of an insulator can be modulated by adjacent *cis*-acting sequences³⁰. Our results reveal that in vertebrates the activity of enhancer boundaries can be controlled by DNA methylation. Some insulators may act not only as fixed boundaries, but also as switches that provide a new kind of modulated gene regulation. □

Methods

Constructions

The 2.5-kb ICR fragment and the 1.6-kb DMD fragment were generated by PCR on genomic DNA with ICRR, 5'-AGGCGCGCCCAAGCTTTGTACAGCGGACCCCAACCTATG, and ICRL, 5'-AGGCGCGCCCAAGCTTTGTCCACCACTTGTCTAAGT, and DMD, 5'-AGGCGCGCCCGTACCTCTGGAGCTCGGACTCCCAATATCA, primers, respectively. The approximate 800-bp HSI fragment was generated with ICRR and HSIR, 5'-GGCGGCCATAGTAGCTATACATTTTCA, and the HS2 fragment was generated with HS2F, 5'-AGGCGCGCTTTATAAGAGGTTGGAACACTTGT, and ICRR. We deleted m1 and m2 from HSI by PCR using the following additional primers: HSIΔm1F, 5'-CCCTATTCTTGGACGTCTGCTGAATCTATTGGAATTCACAAATGGCAATGC; HSIΔm1R, 5'-GATTACAGCAGCTCCAAAGAATAGGG; HSIΔm2F, 5'-GACTCGGACTCCCAATCAACAAGGACGATTGCAACTGATTGAGTTTC; HSIΔm2R, 5'-CCTTGTGATTGGGAGTCCGAGTC; HS2Tm4R, 5'-AGGCGCGCCCAAGCTGAAGGAGCTACCAAGAA; HS2F, 5'-AGGCGCGCTTTATAAGAGGTTGGAACACTTGT; HS2Tm3F, 5'-AGGCGCGCCAGAGAACTGACTCATTCCCTACAC; HS2Δm3F, 5'-AGAAGCTGTTATGTGCAACAAGGAGCGATTTCATCCAGCAATATCC; HS2Δm3R, 5'-CCCTGTGTCACATAACAGCTTCT. The fragments Δm3 and Δm4 are roughly 200-bp truncations of the 5' and 3' ends of HS2 and are generated by PCR with the primer pairs HS2Tm3F/ICRR and HS2F/HS2Tm4R, respectively. In fragment Δm3Δm4, roughly 90 bp spanning the m3 site were internally deleted, the deletion of m4 results from a 3' truncation. This was accomplished by two-step overlapping PCR using the primer pairs HS2Δm3F/HS2Tm4R and HS2F/HS2Δm3R on a DMD clone template. The products of these reactions were gel-purified and mixed, and the final product was amplified by PCR with primer pairs HS2F/HS2Tm4R. Internal deletions of about 90-bp fragments spanning m1 and m2 from HSI were generated by first amplifying with primer pairs HSI1F/HSIΔm1R, HSIΔm1F/HSIR, HSIΔm1F/HSIΔm2R, HSI1F/HSIΔ2R and HSIΔm2F/HSIR. To generate singly or doubly deleted fragments, the products of these reactions were gel-purified, mixed accordingly and amplified with HSI1F/HSIR. The resulting fragments were sub-cloned into pNI⁺ after addition of the appropriate linkers where necessary. For enhancer-blocking assays with m3, 5'-GCTGTTATGTGCAACAAGGGAACGGATGCTACCGCGCGTGGCAGCATACTCTATATATCGTGGCCCAATGCTGCCAACTGGGGAGCGATTTCATTC was directly synthesized with the appropriate restriction sites at its ends and cloned into pNI at either the *Ascl* or *NdeI* sites.

Enhancer-blocking assays

Assays were performed as described⁶. Each reporter carried a gene for neomycin resistance driven by the human γ -globin promoter and a strong enhancer derived from a mouse β -globin LCR element. This was stably transfected into the human erythroleukaemia line K562, and the number of colonies resistant to G418 was counted. As in earlier studies^{6–8}, when the fragment being tested was in the 'up' or 'in' position, a 1.2-kb chicken insulator element was placed 3' of the promoter to block the effects of an enhancer on an adjacent tandemly integrated promoter. The 1.2-kb insert was not present when the test fragment was in the 'down' position, as the purpose was only to measure silencing activity on the adjacent promoter.

DNA-binding assays

Protein preparations and gel mobility-shift assays were performed as described⁶. We generated *in vitro* translated chicken CTCF by transcription/translation of a clone provided by A. West. We incubated 20 μ l reactions at room temperature with 1–2 μ l of protein and 20 fmol of labelled duplex DNA. Probes were annealed with duplexes of the following sequences: m1, AGGCGCGCCGTTGTGGGGTTTATACGCGGGAGTTG CCGCGTGGTGGCAGCAAAATCGATTGCGCAAACTCAAGAGCGCGCGCCT; m2, AGGCGCGCCAAATCCTTTGTGTGTAAAGACCAGGGTTGCCGACGGCGGCAGTGAAGTCTCGTACATCGCAGTCCGCGCGCCT; m3, AGGCGCGCCCTGTATTGTG TGAACAGGGAACGAGTCTACCGCGCGGTGGCAGCATACTCTATATATCGTGGCCCAAGGCGCGCCT; 4m, AGGCGCGCCACGCTGTGCAGATTGGCTATAGCTAAATGACAGACGATGCCGCGTGGTGGCAGTACAATACTACATATGGC GCGCCT; h1, GCCCTGATGGCGCAGAATCGGCTGTACGTGTGGAATCAGAA GTGGCGCGCGCGGCGAGTCAGGCTCACACATCACAGCCGAGCAGCGC; FII, AGGCGCGCCCCAGGATGTAATTACGTCCCTCCCGCTAGGGGGCAGCAG GCGCGCCT; FIIx3', AGGCGCGCCCCAGGATGTAATTACGTCCCTCCCAAT AGCTTTTCAGCAGCGCGCCT; m1x3', AGGCGCGCCTGCTGAATCAGTTGT

GGGGTTTATACGCGGGAGTTGAATATGTTGTTACTCAAATCGATTGCGCCAAAC CCGCGCGCCT. Oligomer modified with 5-methyl C were generated by direct synthesis with the appropriate phosphoramidite (Glen Research) on an ABI synthesizer. Competitions were carried out by the simultaneous addition of a 50-fold excess of unlabelled DNA duplex. We carried out supershifts by adding antibody raised against a conserved carboxyl-terminal CTCF peptide (APNGDLTPMILSMMD) to a standard gel-shift reaction of 2 μ l followed by 2 h incubation at 23 °C.

Received 11 January 2000; accepted 5 April 2000.

- Yoo-Warren, H., Pachnis, V., Ingram, R. S. & Tilghman, S. M. Two regulatory domains flank the mouse *H19* gene. *Mol. Cell. Biol.* **8**, 4707–4715 (1988).
- Bartolomei, M. S., Zemel, S. & Tilghman, S. M. Parental imprinting of the mouse *H19* gene. *Nature* **351**, 153–155 (1991).
- DeChiara, T. M., Robertson, E. J. & Efstratiadis, A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64**, 849–859 (1991).
- Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M. & Bartolomei, M. S. A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. *Nature Genet.* **9**, 407–413 (1995).
- Thorvaldsen, J. L., Duran, K. L. & Bartolomei, M. S. Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes Dev.* **12**, 3693–3702 (1998).
- Bell, A. C., West, A. G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396 (1999).
- Chung, J. H., Whiteley, M. & Felsenfeld, G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**, 505–514 (1993).
- Chung, J. H., Bell, A. C. & Felsenfeld, G. Characterization of the chicken beta-globin insulator. *Proc. Natl Acad. Sci. USA* **94**, 575–580 (1997).
- Tremblay, K. D., Duran, K. L. & Bartolomei, M. S. A 5' 2-kilobase-pair region of the imprinted mouse *H19* gene exhibits exclusive paternal methylation throughout development. *Mol. Cell. Biol.* **17**, 4322–4329 (1997).
- Szabo, P. E., Pfeifer, G. P. & Mann, J. R. Characterization of novel parent-specific epigenetic modifications upstream of the imprinted mouse *H19* gene. *Mol. Cell. Biol.* **18**, 6767–6776 (1998).
- Khosla, S., Aitchison, A., Gregory, R., Allen, N. D. & Feil, R. Parental allele-specific chromatin configuration in a boundary-imprinting-control element upstream of the mouse *H19* gene. *Mol. Cell. Biol.* **19**, 2556–2566 (1999).
- Hark, A. T. & Tilghman, S. M. Chromatin conformation of the *H19* epigenetic mark. *Hum. Mol. Genet.* **7**, 1979–1985 (1998).
- Frevel, M. A., Hornberg, J. J. & Reeve, A. E. A potential imprint control element: identification of a conserved 42 bp sequence upstream of *H19*. *Trends Genet.* **15**, 216–218 (1999).
- Stadnick, M. P. *et al.* Role of a 461-bp G-rich repetitive element in *H19* transgene imprinting. *Dev. Genes. Evol.* **209**, 239–248 (1999).
- Jinno, Y. *et al.* Mouse/human sequence divergence in a region with a paternal-specific methylation imprint at the human *H19* locus. *Hum. Mol. Genet.* **5**, 1155–1161 (1996).
- Frevel, M. A., Sowerby, S. J., Petersen, G. B. & Reeve, A. E. Methylation sequencing analysis refines the region of *H19* epimutation in Wilms tumor. *J. Biol. Chem.* **274**, 29331–29340 (1999).
- Hark, A. T. *et al.* CTCF mediates methylation-sensitive enhancer blocking activity at the *H19/Igf2* locus. *Nature* **405**, 486–489 (2000).
- Webber, A. L., Ingram, R. S., LeVore, J. M. & Tilghman, S. M. Location of enhancers is essential for the imprinting of *H19* and *Igf2* genes. *Nature* **391**, 711–715 (1998).
- Schmidt, J. V., LeVore, J. M. & Tilghman, S. M. Enhancer competition between *H19* and *Igf2* does not mediate their imprinting. *Proc. Natl Acad. Sci. USA* **96**, 9733–9738 (1999).
- Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. & Tilghman, S. M. Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* **375**, 34–39 (1995).
- Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365 (1993).
- Eggenschwiler, J. *et al.* Mouse mutant embryos overexpressing IGF-II exhibit phenotypic features of the Beckwith–Wiedemann and Simpson–Golabi–Behmel syndromes. *Genes Dev.* **11**, 3128–3142 (1997).
- Sun, F. L., Dean, W. L., Kelsey, G., Allen, N. D. & Reik, W. Transactivation of *Igf2* in a mouse model of Beckwith–Wiedemann syndrome. *Nature* **389**, 809–815 (1997).
- Weksberg, R., Shen, D. R., Fei, Y. L., Song, Q. L. & Squire, J. Disruption of insulin-like growth factor 2 imprinting in Beckwith–Wiedemann syndrome. *Nature Genet.* **5**, 143–150 (1993).
- Joyce, J. A. *et al.* Imprinting of *IGF2* and *H19*: lack of reciprocity in sporadic Beckwith–Wiedemann syndrome. *Hum. Mol. Genet.* **6**, 1543–1548 (1997).
- Steinman, M. J. *et al.* Loss of imprinting of *IGF2* is linked to reduced expression and abnormal methylation of *H19* in Wilms' tumour. *Nature Genet.* **7**, 433–439 (1994).
- Okamoto, K., Morison, I. M., Taniguchi, T. & Reeve, A. E. Epigenetic changes at the insulin-like growth factor II/*H19* locus in developing kidney is an early event in Wilms tumorigenesis. *Proc. Natl Acad. Sci. USA* **94**, 5367–5371 (1997).
- Reik, W. *et al.* Imprinting mutations in the Beckwith–Wiedemann syndrome suggested by altered imprinting pattern in the *IGF2-H19* domain. *Hum. Mol. Genet.* **4**, 2379–2385 (1995).
- Taniguchi, T., Sullivan, M. J., Ogawa, O. & Reeve, A. E. Epigenetic changes encompassing the *IGF2/H19* locus associated with relaxation of *IGF2* imprinting and silencing of *H19* in Wilms tumor. *Proc. Natl Acad. Sci. USA* **92**, 2159–2163 (1995).
- Zhou, J. & Levine, M. A novel *cis*-regulatory element, the PTS, mediates an anti-insulator activity in the *Drosophila* embryo. *Cell* **99**, 567–575 (1999).

Acknowledgements

We thank S. Tilghman and members of her laboratory for sharing data before publication, and R. Martin, C. Trainor and members of the Felsenfeld laboratory for discussions and comments on the manuscript.

Correspondence and requests for materials should be addressed to G.F. (e-mail: gary.felsenfeld@nih.gov).

CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus

Amy T. Hark*, Christopher J. Schoenherr*, David J. Katz, Robert S. Ingram, John M. Leverage & Shirley M. Tilghman

Howard Hughes Medical Institute and Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA

* These authors contributed equally to this work

The *Insulin-like growth factor 2* (*Igf2*) and *H19* genes are imprinted, resulting in silencing of the maternal and paternal alleles, respectively. This event is dependent upon an imprinted-control region two kilobases upstream of *H19* (refs 1, 2). On the paternal chromosome this element is methylated and required for the silencing of *H19* (refs 2–4). On the maternal chromosome the region is unmethylated and required for silencing of the *Igf2* gene 90 kilobases upstream². We have proposed that the unmethylated imprinted-control region acts as a chromatin boundary that blocks the interaction of *Igf2* with enhancers that lie 3' of *H19* (refs 5, 6). This enhancer-blocking activity would then be lost when the region was methylated, thereby allowing expression of *Igf2* paternally. Here we show, using transgenic mice and tissue culture, that the unmethylated imprinted-control regions from mouse and human *H19* exhibit enhancer-blocking activity. Furthermore, we show that CTCF, a zinc finger protein implicated in vertebrate boundary function⁷, binds to several sites in the unmethylated imprinted-control region that are essential for enhancer blocking. Consistent with our model, CTCF binding is abolished by DNA methylation. This is the first example, to our knowledge, of a regulated chromatin boundary in vertebrates.

Chromatin boundaries create independent regions of gene expression by preventing the expansion of heterochromatin and by blocking enhancers from activating inappropriate promoters^{8–10}. An element with enhancer-blocking activity must reside between a promoter and an enhancer to exert its repressive effect. The unmethylated *Igf2/H19* imprinted-control region (ICR) lies between the enhancers and the *Igf2* gene on the maternal chromosome, in a position where it could silence *Igf2* by enhancer blocking. Moving the enhancers from their 3' location to a position midway

between the two genes resulted in expression of *Igf2* and silencing of *H19*, as predicted by a boundary model⁵. Further support for this model was produced by deleting a region within the ICR which resulted in the expression of *Igf2* on the maternal chromosome². In addition, the ICR contains two maternal-specific nuclease hypersensitive regions: HS1, which maps between –4.1 and –3.6 kilobases (kb), and HS2, between –2.7 to –2.1 kb upstream of *H19* (refs 6, 11). Both are candidate sites for boundary protein binding.

We tested the boundary model in a mouse transgene assay using the two constructs illustrated in Fig. 1a. The first construct, IHIL, contains a portion of the *H19* ICR and promoter, the *Mus spretus* *H19* gene, and its endodermal enhancers. This is followed by the same ICR and promoter sequences fused to the *luciferase* (*Luc*) gene. Thus, *Luc* is separated from the enhancers by the putative boundary; *H19*, however, has access to the enhancers. The second construct, IHΔIL, is identical to IHIL except for a 670-base-pair (bp) deletion that removes HS2 within the ICR adjacent to *Luc*. If HS2 has blocking activity, its deletion should allow activation of *Luc* in the IHΔIL transgenic lines. Indeed, northern analysis indicated that the average expression of *Luc* in the IHΔIL lines was 14-fold higher than that in the IHIL transgenic lines (Fig. 1c, d). In contrast, we showed that the absolute amounts and the degree of variation of transgenic *H19* RNA were similar between the IHIL and the IHΔIL lines (Fig. 1b). This eliminates differences in *Luc* expression arising from site-of-integration effects. Rather, these results show that a 670-bp region within the ICR has enhancer-blocking activity.

To determine whether DNA methylation affected expression from the IHIL and IHΔIL transgenic lines, we examined the methylation status of the transgene upon both maternal and paternal inheritance. In most lines the transgene was unmethylated and the amounts of *H19* and *Luc* RNAs were similar, irrespective of parental inheritance (data not shown). This finding is unexpected, given the frequency with which multi-copy *H19* transgenes exhibit paternal methylation and imprinted expression^{3,12,13}. The difference may be due to the truncated enhancer fragment in IHIL and IHΔIL mice, or to the presence of luciferase, as both changes have been associated with loss of paternal methylation in transgenes^{12,13}.

Next we analysed the enhancer-blocking activity of the ICR in a stable co-transfection assay in Hep3B cells. In this assay, a reporter gene (*neo*) is controlled by the *H19* promoter and the endodermal enhancers (Fig. 2a). We inserted different test fragments between the promoter and enhancers and assayed for their effects on *neo* RNA expression after co-transfection and selection with a hygromycin resistance gene. A portion of the ICR between –3.8 and

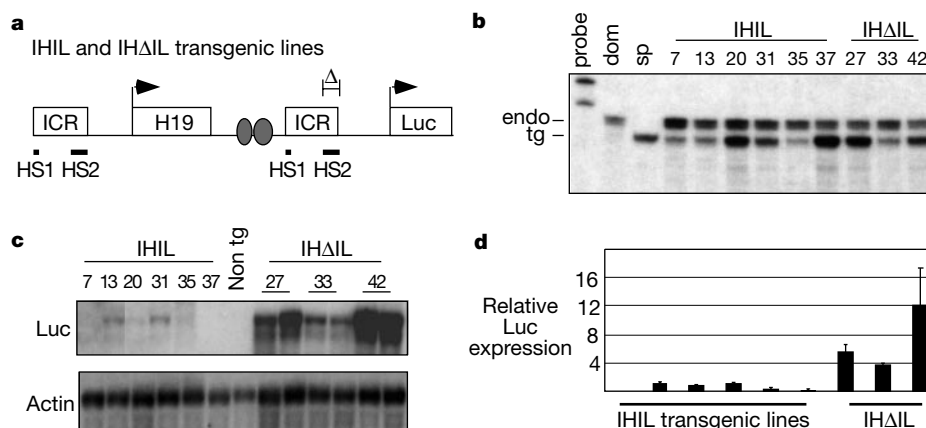


Figure 1 Testing boundary activity in transgenic mice. **a**, Structure of the transgene constructs: endodermal enhancers, two closed ovals; line below Δ , position of a 668-bp deletion of HS2 within the IHΔIL transgene. **b**, Allele-specific *H19* RNase protection assay of neonatal liver from maternal heterozygotes transgenic for IHIL or IHΔIL. Endo, endogenous RNA; tg, transgene RNA; probe, undigested probe; dom, non-transgenic FVB

control; sp, *M. spretus* control. **c**, Northern analysis of RNAs in **b** using *Luc* and *actin* probes. Non Tg, nontransgenic littermate. **d**, Average *Luc* values for IHIL and IHΔIL transgenic lines. *Luc* expression values normalized to *actin* and adjusted relative to line IHIL 31, which was set to 1. Error bars, s.d. of 2–6 individuals.

–1.8 kb (EB in Fig. 2b) inserted into the blocking site reduced the expression of *neo* to around that seen in the absence of the enhancers. In the opposite orientation, EB reduced *neo* expression to a lesser degree (Fig. 2b, EB-Rev). This modest orientation dependence was seen for all fragments with blocking activity (data not shown). To control for distance effects, we used a 2.3-kb *HindIII* fragment of phage λ , which had a minimal effect on *neo* expression (Fig. 2b). Insertion of EB downstream of the enhancers, where its silencing activity can be assessed (EB-Sil), resulted in no decrease in *neo* expression. These results confirm the findings from transgenic mice and indicate that the *H19* ICR has enhancer-blocking activity and is not a silencer.

The hypersensitive regions, HS1 and HS2, of the ICR each displayed partial enhancer-blocking activity, compared to a 2.2-kb *NcoI*–*BglII* fragment (NBg) that includes both hypersensitive regions (Fig. 2b). Multimers of HS1 and HS2 exhibited increases in blocking activity which were dependent on copy number (Fig. 2b). This indicates that both hypersensitive regions of the ICR are required for full enhancer-blocking activity *in vivo*. In contrast, the sequence between HS1 and HS2 (IVS) had minimal blocking activity.

We also tested several fragments within the differentially methylated region upstream of the human *H19* gene in the transfection assay. This region is different in sequence and organization from the mouse gene and is composed of several copies of two different roughly 400-bp direct repeats¹⁴ (A and B in Fig. 3a). When a 1.3-kb *EagI* fragment (Ea) containing 2.5 copies of the B repeat and a 2.5-kb *EagI*–*BglII* fragment that contained 3.5 copies of the B repeat and

one copy of the A repeat were assayed, each exhibited enhancer-blocking activity (Fig. 3c and data not shown). Placing the Ea fragment in the silencing position actually enhanced transcription. Two different B repeats, B1 and B2, exhibited enhancer-blocking activity dependent on copy number.

We then compared the mouse and human ICRs to identify candidate protein-binding sites. A CpG-rich 45-bp sequence in the human B repeats was found to be 60% identical to two sites in HS1 and three sites in HS2, as reported^{14,15} (Fig. 3b). CTCF, a zinc-finger protein implicated in both transcriptional activation and repression^{16–19}, is required for the enhancer-blocking function of several vertebrate boundary or insulator elements⁷. Comparison of the 45-bp consensus sequence to the CTCF binding site in the chicken β -globin insulator (FII) shows a high degree of similarity within a 14-bp core (Fig. 3b). To determine whether the 45-bp sequences bound CTCF, we generated recombinant CTCF and verified that it bound strongly and specifically to the FII binding site (Fig. 4a, lanes 4–7)⁷. CTCF also bound to an oligonucleotide containing the 45-bp conserved sequence within the human B1 repeat (lanes 8–11), but with lower affinity than the FII binding site. In addition, four sites within HS1 and HS2 in the mouse ICR also bound specifically to CTCF *in vitro* (Fig. 4a, lanes 13–17). One site in HS2, predicted by sequence similarity to the other sites (site MS4 in Fig. 3b), failed to bind CTCF *in vitro*. Mutation of the core CpG-rich conserved region within B1 and MS2 (mutB1; Fig. 3b) abolished DNA binding and rendered them ineffective competitors for FII or B1 binding (Fig. 4a and data not shown).

To confirm that CTCF was present in the cells that were used to detect boundary activity, we performed gel-shift assays using a Hep3B nuclear extract (Fig. 4b). FII, B1 and the four mouse CTCF sites bound a protein in the nuclear extract with similar mobility and binding specificity to recombinant CTCF (Fig. 4b, lanes 18–25

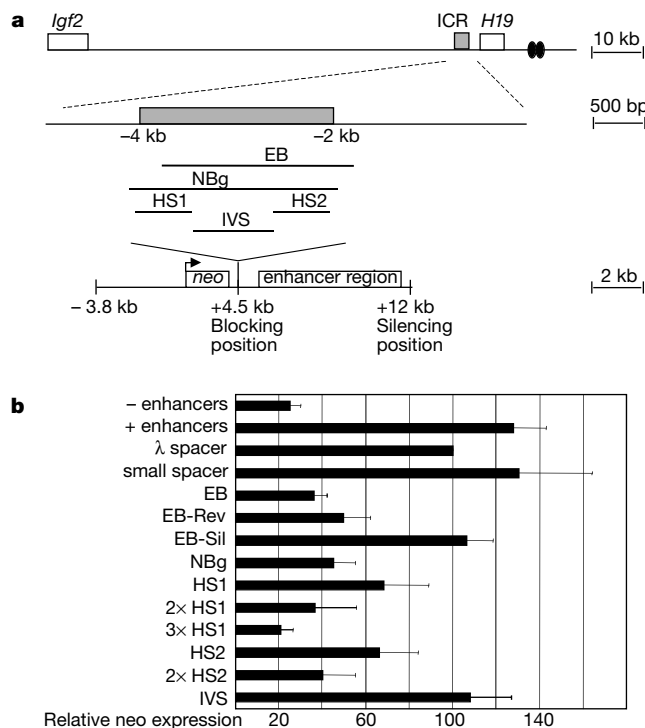


Figure 2 Boundary activity in stable transfection of Hep3B cells. **a**, The enhancer-blocking construct and regions of the mouse *H19* ICR tested for activity. The first line represents the *H19/Igf2* locus, with enhancers shown as ovals. **b**, Relative levels of *neo* RNA following stable transfection into Hep3B cells. The test fragments are defined as: – enhancers, reporter without the enhancer region; λ spacer, a 2.3-kb *HindIII* fragment of phage λ ; small spacer, a 680-bp *BamHI*–*AclI* fragment within the *H19* gene. EB-Rev, EB in the 3' to 5' orientation; EB-Sil, EB in the silencing position. Multimers of HS1 and HS2 are indicated by 2 $\times$ and 3 $\times$. Other designations depicted in **a**. Neo expression levels are the average of 3–6 experiments, normalized to actin RNA and expressed relative to the λ spacer construct, set to 100. Error bars, s.d.

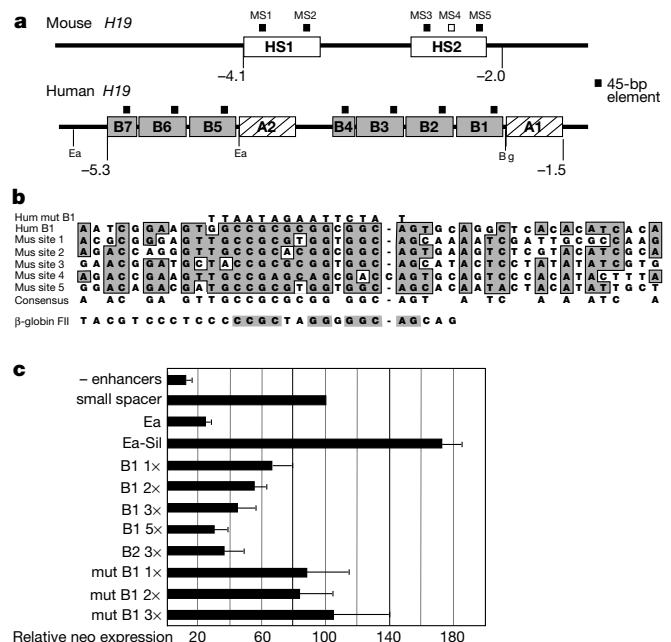


Figure 3 Structure and functional analysis of the human ICR. **a**, Structure of mouse and human ICRs. MS4 (open square) does not bind CTCF. Ea, *EagI*; Bg, *BglII*. **b**, DNA sequence alignment of the 45-bp sequences within the human and mouse ICRs. The bases conserved between the β -globin FII site and the *H19* sequences are highlighted. **c**, Relative *neo* levels, expressed as a percentage of the construct containing the spacer, following stable transfection. Small spacer, as in Fig. 2b; Ea, a 1.3-kb *EagI* fragment containing B5–B7; Ea-Sil, the Ea fragment in the silencing position; B1, the B1 repeat; mutB1, the B1 repeat with the core conserved region mutated as in **b**.

and data not shown). In addition, polyclonal antibodies specific for CTCF inhibited formation of a complex between CTCF and the B1 probe, whereas non-immune IgG had no effect (Fig. 4c, lanes 1–3).

To test directly whether CTCF binding is required for the enhancer-blocking function of the ICR, we compared the blocking activity of multimers of the human B1 repeat with multimers in which the CTCF site had been mutated (mutB1 in Fig. 3b). The mutation that abolishes binding of CTCF (Fig. 4a, b) also eliminates boundary function (Fig. 3c), indicating that this site is required for the activity of B1, and that CTCF is probably an essential mediator of enhancer-blocking activity at *H19*.

One implication of the boundary proposal for imprinting is that CpG methylation of the paternal ICR inhibits the formation of the boundary²⁰. To test whether DNA methylation can directly inhibit CTCF binding, we performed *in vitro* binding assays with probes methylated at all CpG residues. This modification inhibited binding of CTCF to the human B1 site as well as the four sites in the mouse ICR (Fig. 4d, lanes 1–10), indicating that the differential methylation mark could have a direct impact on boundary formation.

We also determined whether CTCF can interact with hemi-methylated DNA using B1 and MS1 probes that were methylated on either strand. When the top strand of B1 and MS1 (Fig. 3b) was methylated, CTCF binding was inhibited, whereas hemi-methylation

of the bottom strand had no effect (Fig. 4c, lanes 11–18). This indicates that the CTCF sites are not symmetric, and that the protein makes important contacts with some of the cytosine residues on the top DNA strand. This result indicates that CTCF may bind to the ICR when DNA is transiently hemi-methylated during replication. As transcription factor binding to DNA has been linked to demethylation^{21,22}, this conditional binding may be a source of epimutations or a mechanism for demethylating the paternal chromosome in the female germline.

These studies support the hypothesis that *Igf2* imprinting requires a chromatin boundary element that lies between the gene and enhancers 3' of the ICR²³. Furthermore, the CTCF protein probably mediates this function *in vivo*, as it binds in a methylation-sensitive manner to sites within the ICR that are required for enhancer blocking. CTCF is a ubiquitously expressed protein¹⁶, consistent with the finding that the nuclease hypersensitivity of HS1 and HS2 is found in all tissues, independent of *H19* expression^{6,11,24}. Methylation-sensitive binding of CTCF provides an explanation for the maintenance of different functions of the maternal and paternal ICRs in somatic tissues. *In vivo* tests will be required to validate the *in vitro* findings and to address CTCF's role in establishing and maintaining the unmethylated state of the maternal chromosome during oogenesis and early development^{6,21,22,25}. □

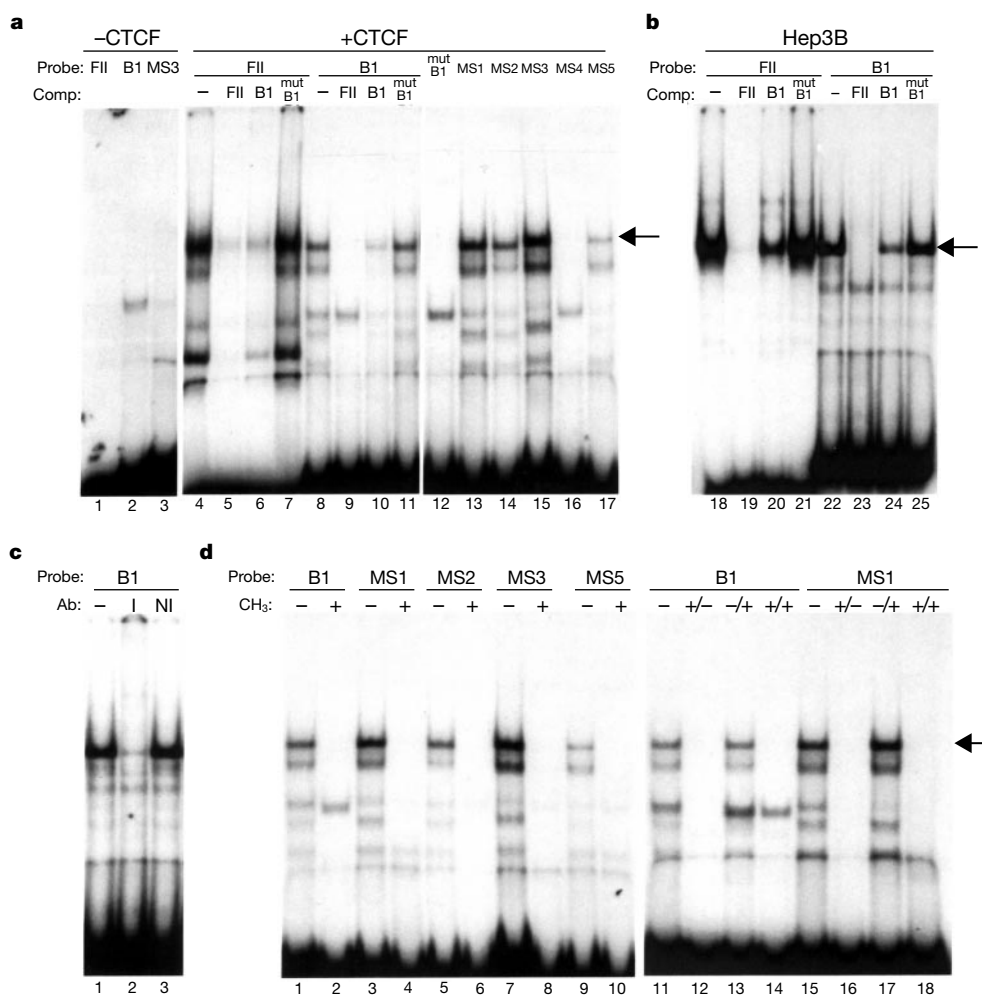


Figure 4 Methylation-sensitive binding of CTCF to multiple sites within the ICR. **a**, Radiolabelled oligonucleotide probes derived from chicken β -globin (Fil), human B1 and mouse ICR (MS1–5, labelled as in Fig. 3a) were incubated in the presence or absence of *in vitro* translated CTCF. Unlabelled oligonucleotides were included as competitors at 100 $\times$ (or 1,000 $\times$ for B1 and mutB1 in lanes 6 and 7) excess. **b**, The same oligonucleotide

probes and competitors incubated with nuclei prepared from Hep3B cells. **c**, Hep3B nuclei incubated with B1 alone, B1 plus anti-CTCF IgG (I) or B1 plus non-immune IgG (NI). **d**, DNA probes were unmethylated (–), hemi-methylated (+/– and –/+) or fully methylated (+, +/+) on all CG residues, and incubated with recombinant CTCF.

Methods

Generation of transgenic mice

The plasmids used to generate the IHIL transgene construct were as described¹³. The *H19* 5' flank and reporter genes were excised and ligated to a 2.4-kb fragment containing the *H19* endodermal enhancers, and the 16-kb product was cloned into Lambda DASHII (Stratagene). A loxP site was inserted at a unique site at the 5' end of the construct. We generated the IHAIL transgene in a similar manner, except that a 668-bp deletion from -2070 to -2739 bp with respect to the *H19* start site was introduced into the 5' flank upstream of *Luc*. The phage DNAs were digested with *NotI* and injected into fertilized FVB eggs. All transgenic animals were bred to FVB.

DNA analysis

We identified founder transgenic mice by Southern blotting. Methylation of the transgenic ICR was assessed using the methylation-sensitive enzymes *HpaII*, *AclI* and *BstUI*; the *H19* and *Luc* promoters were assayed with *HpaII*.

RNA analysis

Total RNA was prepared from postnatal day 2 liver of transgenic animals by extraction with LiCl-urea²⁶. The *H19* allele-specific RNase protection assay was performed essentially as described, using the Ambion RPA II kit and linearizing the *H19* template with *HindIII*²⁷. Northern analysis was performed with DNA probes from the *luciferase*, β -actin and neomycin resistance genes. Quantification was performed using a Molecular Dynamics Phosphorimager.

Stable transfection of Hep3B cells

The test construct was made by fusing the *EcoRI*-*DraIII* fragment of mouse *H19* (-3.8 kb to +10 bp) to a neomycin cassette in pBluescript II. A 3' *SmaI*-*BamHI* fragment of *H19* (+4.5 to +12 kb) containing the enhancers was inserted downstream of the *neo* gene. Test fragments were inserted into either a *XhoI* (blocking position) or *NotI* (silencing position) site. Plasmids were prepared using a Maxiprep Kit (Qiagen) and linearized with *Sall* before transfection. Equimolar (0.5 pmol) amounts of each construct were electroporated (BioRad Gene Pulser) in 2×10^6 Hep3B cells along with 1 μ g of a plasmid encoding a hygromycin resistance gene linearized with *NotI*. Cells were plated onto 100-mm plates and selection with hygromycin (100 μ g ml⁻¹) was initiated 60 h later. Cells were passaged after 10 days, harvested at confluence and RNA was prepared using Qiashredder and RNeasy kits (Qiagen).

In vitro DNA binding

Full length mouse CTCF complementary DNA was isolated by reverse transcription (RT)-PCR, cloned into pCRII (Invitrogen) and transferred into the *EcoRI* site of pcDNA3.1/*HISB* (Invitrogen) forming an in-frame fusion protein. Capped messenger RNA from this construct was used to program a rabbit reticulocyte lysate system according to the manufacturer's protocols (Ambion). One microlitre of the lysate was added to a binding buffer containing 20 mM HEPES (pH 7.5), 100 mM KCl, 2.5 mM MgCl₂, 5% glycerol, 3 μ g ml⁻¹ single strand salmon sperm DNA and 125 μ g ml⁻¹ poly dI-dC and incubated for 30 min on ice. Probes and competitors were added simultaneously to the reaction and incubated for 20 min at room temperature before electrophoresis on a 4% polyacrylamide gel in 0.25 $\times$ TBE. Binding reactions using roughly 10^6 Hep3B (or 10^5 using the FII probe) nuclei²⁸ were performed similarly except that the binding buffer contained 250 mM KCl, 3% glycerol, 18 μ g ml⁻¹ single-strand salmon sperm DNA and 0.1% NP-40. Antibody inhibition experiments were performed by adding 2 μ l of anti-CTCF IgG (Upstate Biotechnology) or non-immune rabbit IgG to Hep3B nuclei in binding buffer and processed as above. The top strands of each oligonucleotide probe were: B1-5' tcgaTTGGTTGTAGTTGTGGAATCGGAAGTGGCGCGCGGCGGAGTGCAGGCTCAC ACATCACAGCCCCGAGCCCGC; mutB1-tcgaTTGGTTGTAGTTGTGGAATCGGAAG TTTAATAGAATTCTATGTGTCAGGCTCACACATCACAGCCCCGAGCCCGC; MS1-tcgaCAGTTGTGGGGTTTATACGCGGGAGTTGCCGCGTGGTGGCAGCAAAATCGA TTGCGCCAAACCTAAAGAG; MS2-tcgaAATCCTTTGTGTGTAAGACACAGGGTT GCCGCACGGCGGCGAGTGAAGTCTCGTACATCGCAGTCTCTAAAC; MS3-aggcgcgcCTGTTATGTGCAACAGGGAACGATGCTACCGCGCGGTGGCAGCATA CTCCTATATCGTGGGCCCAAGgcgcct; MS4-tcgaTCAGTGGCTGGTAAGACCGAA GTTGGCGAGCAGCGACAGTGCAGTCCCATCATCTTTATCATAGAGGTG; MS5-tcgaTTGGCTATAGCTAAATGGACAGACGATGCCGCGTGGTGGCAGTACAATACTAC ATATTGCTCGGAGACG; and FII-aggcgcgCCCCAGGGATGTAATTACGTCCCT CCCCCGTAGGGGGGACGAGgcgcct. The uppercase sequences represent actual genomic sequence, and lowercase sequences were used for cloning. Methylated probes were identical to their unmethylated counterparts except for the absence of cloning sequences.

Received 11 January; accepted 5 April 2000.

- Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. & Tilghman, S. M. Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* **375**, 34–39 (1995).
- Thorvaldson, J. L., Duran, K. L. & Bartolomei, M. S. Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H10* and *Igf2*. *Genes Dev.* **12**, 3693–3702 (1998).
- Bartolomei, M. S., Webber, A. L., Brunkow, M. E. & Tilghman, S. M. Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev.* **7**, 1663–1673 (1993).
- Ferguson-Smith, A. C., Sasaki, H., Cattanaach, B. M. & Surani, M. A. Parental-origin-specific epigenetic modifications of the mouse *H19* gene. *Nature* **362**, 751–755 (1993).
- Webber, A., Ingram, R. L., Levorse, J. & Tilghman, S. M. Location of enhancers is essential for imprinting of *H19* and *Igf2*. *Nature* **391**, 711–715 (1998).
- Hark, A. T. & Tilghman, S. M. Chromatin conformation of the *H19* epigenetic mark. *Hum. Mol. Genet.* **7**, 1979–1985 (1998).
- Bell, A. C., West, A. G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396 (1999).
- Kellum, R. & Schedl, P. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**, 941–950 (1991).
- Kellum, R. & Schedl, P. A group of *scs* elements function as domain boundaries in an enhancer-blocking assay. *Mol. Cell. Biol.* **12**, 2424–2431 (1992).
- Kellum, R. & Elgin, S. C. R. Chromatin boundaries: punctuating the genome. *Curr. Biol.* **8**, R521–R524 (1998).
- Khosla, S., Aitchison, A., Gregory, R., Allen, N. D. & Fell, R. Parental allele-specific chromatin configuration in a boundary-imprinting-control element upstream of the mouse *H19* gene. *Mol. Cell. Biol.* **19**, 2556–2566 (1999).
- Elson, D. A. & Bartolomei, M. S. A 5' differentially methylated sequence and the 3' flanking region are necessary for *H19* transgene imprinting. *Mol. Cell. Biol.* **17**, 309–317 (1997).
- Pfeifer, K., Leighton, P. A. & Tilghman, S. M. The structural *H19* gene is required for its own imprinting. *Proc. Natl Acad. Sci. USA* **93**, 13876–13883 (1996).
- Frevel, M. A., Hornberg, J. J. & Reeve, A. E. A potential imprint control element: identification of a conserved 42 bp sequence upstream of *H19*. *Trends Genet.* **15**, 216–218 (1999).
- Stadnick, M. P. *et al.* Role of a 461-bp G-rich repetitive element in *H19* transgene imprinting. *Dev. Genes Evol.* **209**, 239–248 (1999).
- Filippova, G. N. *et al.* An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol. Cell. Biol.* **16**, 2802–2813 (1996).
- Vostrov, A. A. & Quitschke, W. W. The zinc finger protein CTCF binds to the APB β domain of the amyloid β -protein precursor promoter. Evidence for a role in transcriptional activation. *J. Biol. Chem.* **272**, 33353–33359 (1997).
- Burcin, M. *et al.* Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone-receptors, is identical to the multivalent zinc finger repressor CTCF. *Mol. Cell. Biol.* **17**, 1281–1288 (1997).
- Awad, T. A. *et al.* Negative transcriptional regulation mediated by thyroid hormone response element 144 requires binding of the multivalent factor CTCF to a novel target DNA sequence. *J. Biol. Chem.* **274**, 27092–27098 (1999).
- Bird, A. P. & Wolffe, A. P. Methylation-induced repression—belts, braces, and chromatin. *Cell* **99**, 451–454 (1999).
- Macleod, D., Charlton, J., Mullins, J. & Bird, A. P. Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* **8**, 2282–2292 (1994).
- Brandeis, M. *et al.* Sp1 elements protect a CpG island from *de novo* methylation. *Nature* **371**, 435–438 (1994).
- Bell, A. C. & Felsenfeld, G. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482–485 (2000).
- Szabo, P. E., Pfeifer, G. P. & Mann, J. R. Characterization of novel parent-specific epigenetic modifications upstream of the imprinted mouse *H19* gene. *Mol. Cell. Biol.* **18**, 6767–6776 (1998).
- Feil, R. & Khosla, S. Genomic imprinting in mammals: an interplay between chromatin and DNA methylation? *Trends Genet.* **15**, 431–435 (1999).
- Auffray, C. & Rougeon, F. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. *Eur. J. Biochem.* **107**, 303–314 (1980).
- Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L. & Tilghman, S. M. An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev.* **9**, 2079–2089 (1995).
- Hagenbuchle, O. & Wellauer, P. K. A rapid method for the isolation of DNA-binding proteins from purified nuclei of tissues and cells in culture. *Nucleic Acids Res.* **20**, 3555–2559 (1992).

Acknowledgements

We thank A. Bell and G. Felsenfeld for communicating results before publication. We also thank J. Shen for her assistance with the sequence alignment analysis and members of our laboratory for helpful discussions. This work was supported by a grant from the National Institute of General Medical Sciences. C.J.S. was a recipient of an NRSA Postdoctoral Fellowship from the NIH and S.M.T. is an Investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be address to S.M.T.

Liquid refreshment

Top tips, filter tips, ergonomic tips and other takes on liquid handling.

Status pipettors

Philip Harris Scientific

www.philipharris.co.uk/scientific

Removable tip ejector

The Status range of variable-volume single- and multi-channel pipettors are compatible with Media pipette tips. The pipettors have a tip ejector positioned to avoid unintentional tip ejection which can be removed for work in deep, narrow tubes. There are seven variable volume pipettors with volume ranges 0.2 µl to 2.0 µl through to 100 µl and 5,000 µl.

Reader Service No. 100

Biochem VacuuCentre

ThermoQuest

www.thermoquest.com

High aspirations

Designed for drawing off surplus liquid from microtitre plates, Petri dishes or cell culture flasks, the VacuuCentre can be used with variously sized pipettes and tips. Handling is claimed to be precise and safe, with a high-volume collecting flask and powerful vacuum pump, allowing simultaneous operation of several workstations. The VacuuCentre is suitable for fluid aspiration in chemistry, biochemistry, biology and medicine.

Reader Service No. 101

Calibrex delivery jet

Socorex

www.socorex.ch

Extension tubing

This system, an accessory to the Calibrex bottle-top dispenser, consists of extension tubing and a delivery jet pen, and allows dispensing within a 60 cm area. The system is screwed

onto the dispenser in replacement of the standard delivery jet. Made of PTFE material, the system is easy to clean and can be autoclaved together with the instrument, or separately.

Reader Service No. 102

Tipster

Bellco Glass

www.bellcoglass.com

Automatic tip-loading

This device automatically loads six templates of tips per container, which are then transferred to tip holders using the pipette tip pickers. The unit comes with a peg and grommet attachment to fit Bellco orbital shakers and accommodates most brands of tips. Suggested applications include cell culture, genetic, chemical and medical laboratories that perform blood serum tests, and any laboratory needing plastic tips to perform tests and protocols. The unit includes the Tipster box, six templates, a collection tray, a template holder, a picker assembly and an anti-skid frame for temporary or permanent mounting.

Reader Service No. 103

Hydra-PP

Robbins Scientific

www.robsci.com

Intelligent plate handling

The Hydra dispenser has a microscope stage that moves source and target plates under the dispensing head with a positioning accuracy of five microns. It also performs automatic syringe washing between samples. The system can be used with all microplates, including deep-well, 384 and 1536 well plates. Functions are programmed using Windows 95/98 or NT. The Hydra can also be used as a stand-alone workstation with plate placement performed by technicians, or combined with the Twister-H or other robotic plate handling systems to form an automated workstation.

Reader Service No. 104

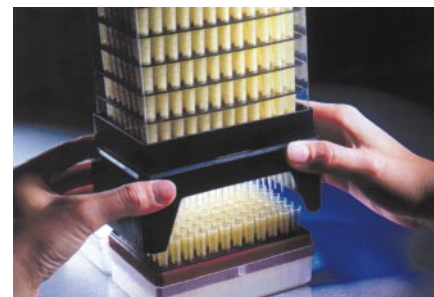
Impact Equalizer

Matrix Technologies www.matrixtechcorp.com

Expandable tip-to-tip spacing

The hand-held Equalizer multichannel pipettor is said to speed up transfer between different types of vessels such as 96, 48 and 24 well plates, 12, 13 and 16 mm tubes, microcentrifuge and PCR tubes. Spacings can be adjusted from 9 mm to 14.15 mm in increments of 0.1 mm using a slide rod. The pipettor has five speed settings, including two ultra-slow speeds for dispensing into electrophoresis gels, and can store up to six 40-step programs. A calibration facility is provided.

Reader Service No. 105



Down they go: Corning's tip loader, post-click.

Costar tip-loading system

Corning Costar

www.corning.com

Point and click

Combining the convenience of pre-racked tips with the economy of bulk tips, this tip-loading device has a 10-layer magazine tower holding 96 tips per layer. The loader and magazine are positioned over an empty rack and the tip loader engaged. The system is pre-filled with RNase- and DNase-free Costar universal fit tips, and is said to occupy 65% less storage space than tips stored in conventional racks. An autoclavable hinged tip rack helps to minimize plastic waste.

Reader Service No. 106

Luckhams super filter tips

Life Sciences International

www.llsi.org

One size fits all

These low-retention filter tips have a pliable top opening, allowing use with many different pipettes and reducing air leaks. They include a hydrophobic filter to prevent aerosol containment and aqueous liquid bypass in the event of over-pipetting. By eliminating the need to purchase pipette-specific tips, inventory control and cost savings are possible. The tips are available in an autoclavable non-sterile version or pre-sterilised using γ irradiation.

Reader Service No. 107

Multiprobe II

Packard Instruments

www.packardinst.com

Multi-tipped, low-volume pipetting

Automated PCR set-up and DNA purification are offered by the Multiprobe II. It provides multi-tipped, low-volume pipetting with a delivery range down to 1 µl. There is a choice of fixed washable or disposable tips for contamination control. The unit can be integrated with a vacuum manifold thermal cycler for automated assaying from sample preparation to reaction set-up and amplification using WinPREP software.

Reader Service No. 108



Spring into action: the Calibrex delivery jet.

new on the market

ViscoPet

Biohit

www.biohit.com

Get into the thick of it

This electronic pipettor was developed to cope with viscous liquids. It is a variant of the BPE 1200, and can be used with long-reach Visco-Tip Capillary tips. The Visco-Tip Capillaries and the option to use Safe Cone Filters in the tip cone make the ViscoPet particularly suitable for applications such as bacteriological food analysis. The untapered shape of the Visco-Tip facilitates pipetting of thick liquids without the risk of tip blockage.

Reader Service No. 109

Millimatic

TriContinent

www.tricontinent.com

Disposable dispenser

From TriContinent, a subsidiary of Hitachi, Chemical, the Millimatic bottle-top microplate filler and washer has an eight-port manifold and can dispense volumes of 0.25–3.0 ml. Although priced to be disposable, the manifold is strong enough for repeated use. It is autoclavable to 120 °C.

Reader Service No. 110



Room at the top: Millimatic poised for action.

LiteTouch

Rainin

www.rainin.com

Ease the strain

Traditional conical shafts and tips seal over a large surface area, making it difficult to determine exactly when a seal is achieved. Rainin pipette tips are cylindrical to create a strong seal over small area, and have a positive-stop 'shelf' that indicates when a seal has been made. The pipette and tip system reduces tip ejection forces, easing removal of the tip from the pipette shaft, and should be particularly suitable for applications involving repeated multiple pipetting steps. The tips are designed for use with the EDP3 electronic pipette.

Reader Service No. 111

SciClone

Zymark

www.zymark.com

Pipetting workstation

This workstation can transfer liquids, add reagents and reformat plates with 96 well and 384 well formats. It comes with an integrated nine-position work-table to accommodate any configuration of microplates, reagents, accessories or disposable pipette tips. The SciClone can be configured with the Auto-Stack system, and comes with software for modification of transfer parameters. Optional extras include a disposable tip server system, an automated dispenser, and a reagent dispenser module that provides bulk reagent dispensing directly onto plates or reservoirs.

Reader Service No. 112

Research Pro

Eppendorf

www.eppendorf.com

Battery powered and rechargeable

Similar to mechanical pipettes in use, this battery-driven pipette can perform reverse pipetting and serial dilution in addition to basic pipetting and dispensing. Users can choose from five pre-set volumes, or program their own pipetting sequences. The pipette is available in a single-channel version for the



Batteries included: the Research Pro pipette.

0.5–5,000 µl range and in 8 12 channels for 0.5–1,200 µl. The capacity is 3,700 strokes at top speed. The area susceptible to aerosol formation can be disassembled and is fully autoclavable.

Reader Service No. 113

765 Dosimat

Metrohm

www.metrohm.co.uk

Push-button dispensing

Suitable for titrating, dosing, dispensing, diluting, pumping or preparing standards, this push-button unit dispenses reagent as required. Dosing and filling speeds can be set manually with a stepless rotary knob or digitally, and filling carried out automatically when the cylinder volume has been reached or by push-button. The unit is suitable for use as a chemical-resistant 'precision pump' with most aggressive and/or toxic agents. It has a data interface RS 232C for connecting a balance, printer or computer.

Reader Service No. 114

These notes are compiled in the Nature office from information provided by the manufacturers.

ADVERTISEMENTS

World's First Universal Quantum Computation Simulator QUANTUM COMPUTER SIMULATOR

for Windows/Macintosh/UNIX

The DEMO version is freely downloadable at our Web site. Many sample algorithms including Shor's factorization algorithm and Grover's database search algorithm which are most popular in the field of quantum computer study and manuals are also available with free of charge.

<http://www.senko-corp.co.jp/qcs>

SENKO
CORPORATION

9-30, Kitasaiwai 2-chome, Nishi-ku,
Yokohama, 220-0004 Japan
PHONE: 81-45-320-2711
FAX: 81-45-320-9342
e-mail: senko@246.ne.jp

BACs • YACs • PACs cDNAs

Research Genetics provides clones and colony membranes for a number of human, mouse and rat libraries as well as a custom screening service for the libraries.

Research Genetics

U.S. or Canada 800-533-4363

U.K. 0-800-89-1393

FAX 256-536-9016

Check out our homepage at
<http://www.resgen.com>

Look no further
for the best value in
Oligo Synthesis

Competitive pricing
Routine 24 hour dispatch • Every Oligo QC verified by PAGE
Rigorous batch QC by PCR and Mass Spec
Web ordering and tracking • Local Freephone numbers
A great deal in any language!!

SIGMA (+44) (0) 1223 839000 • fax (+44) (0) 1223 839200
GENOSYS email: info@sigma-genosys.co.uk • www.sigma-genosys.co.uk

READER ENQUIRY NO. 44

READER ENQUIRY NO. 33